AOC 10-7

POPQAE 50 (1) 1–142 (1986)
ISSN 0033–362X

Public Opinion Quarterly

SYMPOSIUM ON POLLS: Is There a Crisis of Confidence?
  Rating the Polls: The Views of Media Elites and the General Public *Andrew Kohut*
  Evaluating Polls with Poll Data *Burns W. Roper*
  Trends in Attitudes Toward and Participation in Survey Research *Stephen Schleifer*
  Surveys on Surveys: Limitations and Potentialities *John Goyder*
Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s *Robert Y. Shapiro and Harpreet Mahajan*
Preelection Polls in Israel: Structural Constraints on Accuracy *Jacob Shamir*
Mass Communication and Political Socialization: Specifying the Effects *Gina M. Garramone and Charles K. Atkin*
Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry *Hans-J. Hippler and Norbert Schwarz*
Interviewers' Voices and Refusal Rates in Telephone Surveys *Lois Oksenberg, Lerita Coleman, and Charles F. Cannell*
A Sample Size Formula for Multiple Regression Studies *Sande Milton*
THE POLLS: ABM and Star Wars *Thomas W. Graham and Bernard M. Kramer*

IN MEMORIAM • NEWS AND NOTES • BOOK REVIEWS

SPRING 1986

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

LEO BOGART — BURNS W. ROPER
NORMAN M. BRADBURN — BERNARD ROSHCO
ANDREW KOHUT AND ELEANOR SINGER, COUNCILLORS-AT-LARGE

Editor: ELEANOR SINGER — *Managing Editor:* INA C. JAEGER

Department Editors:
PHILIP A. HARDING, *News and Notes*
MARC B. GLASSMAN, *Book Reviews*
PHYLLIS M. ENDRENY, *Book Notes*
TOM W. SMITH, *The Polls*

EDITORIAL BOARD

JAMES R. BENIGER
JOAN S. BLACK
NORMAN M. BRADBURN
CHARLES F. CANNELL
MURIEL G. CANTOR
ALBERT H. CANTRIL
DONALD DILLMAN
MARTIN R. FRANKEL
HERBERT J. GANS
ALBERT E. GOLLIN
ROBERT HORNIK
MARY R. JACKMAN
F. THOMAS JUSTER
RONALD C. KESSLER
MARK R. LEVY
PHILIP MEYER
J. RONALD MILAVSKY
WARREN J. MITOFSKY
ELISABETH NOELLE-NEUMANN
STANLEY PRESSER
DONALD F. ROBERTS
MORRIS ROSENBERG
CHARLES R. WRIGHT
DANIEL YANKELOVICH

Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. and at additional mailing offices. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student-subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 SPRING 1986 NUMBER 1

Is There a Crisis of Confidence?

Rating the Polls: The Views of Media Elites and the General Public *Andrew Kohut* 1

Evaluating Polls with Poll Data *Burns W. Roper* 10

Trends in Attitudes Toward and Participation in Survey Research *Stephen Schleifer* 17

Surveys on Surveys: Limitations and Potentialities *John Goyder* 27

Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s *Robert Y. Shapiro and Harpreet Mahajan* 42

Preelection Polls in Israel: Structural Constraints on Accuracy *Jacob Shamir* 62

Mass Communication and Political Socialization: Specifying the Effects *Gina M. Garramone and Charles K. Atkin* 76

Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry *Hans-J. Hippler and Norbert Schwarz* 87

Interviewers' Voices and Refusal Rates in Telephone Surveys *Lois Oksenberg, Lerita Coleman and Charles F. Cannell* 97

A Sample Size Formula for Multiple Regression Studies *Sande Milton* 112

IN MEMORIAM: Herbert H. Hyman, 1918–1985 *Paul B. Sheatsley* 119

NEWS AND NOTES *Luis C. Torres-Bohl, Guest Editor* 121

The Polls: ABM and Star Wars: Attitudes Toward Nuclear Defense, 1945–1985
Thomas W. Graham and Bernard M. Kramer 125

Book Reviews
Marc B. Glassman, Editor 135

Herbert McClosky and John Zaller, *The American Ethos: Public Attitudes Toward Capitalism and Democracy*
Lee Sigelman 135

Rita J. Simon, *Public Opinion and the Immigrant, Print Media Coverage, 1880–1980* *Ronald Humphrey* 137

John P. Murray and Gavriel Solomon, Eds., *The Future of Children's Television* *Steven T. McDermott* 138

Book Notes *Phyllis M. Endreny, Editor* 140

NOTICE TO CONTRIBUTORS

Three copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic-references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

Public Opinion Quarterly is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

Public Opinion Quarterly has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

Is There a Crisis of Confidence?

For the May 17, 1985 plenary session of the AAPOR Annual Conference, Andrew Kohut and Burns W. Roper were invited by Conference Chair Eleanor Singer to address a question often encountered during the preceding year: Is there a crisis of confidence? Both generously responded to the invitation by fielding special questions bearing on that topic in the Spring of 1985. Their reports, first presented at the 1985 Conference, appear below. Two other papers given at the 1985 Conference bear very directly on the "crisis of confidence" theme, and they are also included in this special section of the *Quarterly*. One, by Stephen Schleifer, shows very clear convergences with Roper's data and conclusions. The other, by John Goyder, raises the question, also acknowledged by the other writers, of how it is possible to use survey research to measure people's attitudes to surveys.

Rating the Polls: The Views of Media Elites and the General Public

ANDREW KOHUT

The credibility of opinion polls was very much at issue during the 1984 campaign. Throughout the year, pollsters and news organizations that conduct polls were subject to a steady stream of criticism about their findings and methods. In a major journalistic treatment of polls in 1984, Barry Siegel of the *Los Angeles Times* wrote: "These examples illustrate something most polling experts know but don't always acknowledge. It is remarkably difficult to measure public opinion accurately" (Siegel, 1984).

As late as five days before one of the most lopsided elections in history, the *Washington Post* wrote: "In these last days before the

---

Abstract To measure the credibility of the polls in 1984, Gallup conducted two surveys, one of the press and one of the general public. Results of the press survey indicated a favorable reaction to the accuracy of the preelection polling in 1984, although there was a division of opinion about whether the polls were better at explaining the vote than predicting it, and whether the polls enhanced or interfered with the electoral process. The public's perception of the polls was a generally positive one both as to their accuracy and their overall effect.

Andrew Kohut is President of The Gallup Organization.

Public Opinion Quarterly Vol. 50:1–9 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-1/$2.50

election, Democratic presidential nominee Walter F. Mondale has hoisted a new banner: 'The polls are dead wrong!' In one sense, he is dead right. Throughout this political year, the variation in poll results has been so pronounced as to call into question the scientific precision of the popular art of opinion-taking" (Haywood, 1984).

While there was variability in the polls, the root cause for lack of confidence in the polls in 1984 lay in their failure to detect Reagan's big victory in 1980. For more than two decades most of the national polls had produced final figures that were either right on target or well within sampling error; 1980 came as a big shock to the public as well as to the pollsters themselves. The polls had failed to live up to expectations, and as a consequence, in 1984 the level of expectations was much lower; many expected the polls to perform badly again.

To some extent this expectation may have been intensified by wishful thinking on the part of Mondale supporters. There was a great deal of discontent with the intrusiveness of polls in the electoral process. As the *Washington Post* quote above attests, in the final months Mondale was certainly running against his deficit in the polls as well as against Reagan.

How well the preelection polls do is significant not only for the small yet growing number of organizations that do public preelection polling but also for the survey research profession at large. The national preelection polls are the most visible expression of the validity of the opinion survey method. If the polls had been as far off the mark in 1984 as they were in 1980, the impact on polling organizations and, more generally, on the survey research profession would have been very severe.

As it turned out, the performance of the polls in 1984 was variable. Some came quite close to the actual division of the vote while others greatly exaggerated the Reagan victory and a few underestimated it. Nonetheless, postelection news coverage gave the national polls generally good reviews, putting less emphasis on their variability and more on the fact that unlike the polls of 1980, they mostly all showed a big Reagan victory in their final result (Claffey, 1984; Rosenbaum, 1984).

Given the extent to which the credibility of the polls was at issue over the course of the 1984 election, a Gallup effort was mounted on behalf of the 1985 AAPOR annual conference to measure reaction to the performance of the polls subsequent to the national election. Two Gallup surveys were conducted—one, a survey of the press (the principal client of national polling organizations); the other, a small survey of the general public, the ultimate audience for published opinion polls. The overriding objective of these surveys was a postelection assessment of image. More specifically, the studies sought to determine what

Table 1. Meeting Expectations

Many news organizations made use of preelection polls in their 1984 election coverage. For each of the pairs of opposites listed below, please indicate your opinion of preelection polls in 1984. (If you feel the item [at the top] completely describes preelection polls, so indicate by circling "1." If you feel the item [below] completely describes preelection polls, so indicate by circling "4." You may, of course, indicate you opinion by circling any number between 1 and 4.)

	All Respondents %		*Executive Managing Editor* %	*Other Editors, Reporters* %
Did a better job than I expected				
1	10	68	5	17
2	58		61	53
3	20	26	23	16
4	6		3	8
Did not do as good a job as I expected				
No answer	6		8	6
	100		100	100
N responses	(246)		(65)	(113)

sort of impression was left on the public and how the polls were judged by members of the news media, who presumably have more informed and experienced viewpoints of preelection surveys.

The Gallup survey of the press consisted of 246 questionnaires administered to members of the news media.[1] The poll was conducted by mail in January 1985. The survey, which was part of an ongoing Gallup program, covered a sampling of senior level editors and reporters as well as publishers and management executives of national and major local market news organizations. Of the 246 questionnaires, 137 were completed by print journalists and 104 by broadcast journalists. The majority, 190, represented editors, broadcast producers, and reporters; 56 were publishers or station managers.

Overall, the results of the press survey indicated a favorable reaction to preelection polling in 1984. First, the performance of the published polls exceeded the expectations of news media respondents (Table 1). The style of question was semantic differential, and most—68 percent—chose the categories nearest the better-than-expected alternative. There was little variation by characteristics or attitudes of respon-

---

[1] The initial mailing was made to 415 potential respondents and followed up with one mail prompt to encourage participation.

Table 2. The Accuracy of Polls

	All Respondents %		Type of Media Print %	Type of Media Broadcast %
Accurate				
1	33	83	29	$\boxed{39}$
2	50		54	45
3	13	14	15	10
4	1		1	2
Inaccurate				
No answer	3		1	4
	100		100	100
N responses	(246)		(137)	(104)

Table 3. Distinctions Among National Polls

%	All Respondents		Type of Media Print %	Type of Media Broadcast %
All national polling organizations did about the same quality job				
1	11	30	7	15
2	19		19	19
3	30	68	29	32
4	38		$\boxed{43}$	32
Some national polling organizations did a better job				
No answer	2		2	2
	100		100	100
N responses	(246)		(137)	(104)

dent, except that senior reporters and middle-level editors were slightly more emphatic in this opinion than their superiors.

Most saw the preelection polls as accurate (Table 2). While this was the overall tenor of opinion, it was somewhat more strongly expressed by media respondents who cover *politics*, and by broadcast-affiliated respondents, especially those who work in TV. The perceived accuracy of polls notwithstanding, most respondents felt that some national organizations did a better job than others. Print journalists expressed this view somewhat more strongly than broadcast journalists (Table 3).

Table 4. Forecasting vs. Explaining

	All Respondents %		Audience	
			National %	Local %
Better at explaining why people vote than forecasting elections				
1	14	43	10	15
2	29		20	31
3	25	54	$\boxed{21}$	26
4	29		$\boxed{41}$	27
Better at forecasting elections than at explaining why people vote				
No answer	3		8	1
	100		100	100
N responses	(246)		(49)	(193)

Table 5. Effect on Process

	All Respondents %		Voted for:	
			Reagan %	Mondale %
Interfere with election process				
1	18	53	11	$\boxed{26}$
2	35		31	$\boxed{36}$
3	35	45	41	29
4	10		15	6
Enhance the election process				
No answer	2		2	3
N responses	(246)		(120)	(102)

Surprisingly, there was a close division of opinions as to whether the polls were better at explaining the vote or better at forecasting it, with the balance of opinion on the side of the latter (Table 4). Those employed by media with a national audience were more inclined to see polls as doing a better job of forecasting.

However, the sharpest and most troubling division of opinion was over whether polls *enhanced* or *interfered* with the electoral process (Table 5). Just about equal numbers of respondents took either side of this issue. On this measure there were some very significant variations

Table 6. Exposure to Public Opinion Polls

Do you follow the results of any public opinion poll regularly in any newspaper or magazine?

	1944 %	*1985* %
Regularly	9	$\boxed{25}$
Occasionally	19	16
No	28	59
Never heard of	44	NA
	100	100
N responses	(2000)	(1023)

in opinion. The following were more apt to take the view that polls *enhance* the process—*publishers, station managers,* broadcasters, and those who work in local markets. Feeling that polls interfere with the process was greatest among respondents from the *West,* those who write for *print,* national media, and senior level executives. However, all of these relationships aside, the single greatest correlate of opinion was partisanship—whose who voted for Mondale were far more inclined to see polls as interfering with the process than were those who voted for Reagan.

The survey of the general public was a limited effort. In March 1985, Gallup administered to a national telephone sample of 1120 adult men and women a series of questions about opinion polls which were first addressed to a national sample in 1944. The historic survey was conducted by Hadley Cantril's Princeton office of Public Opinion Research and initially reported by Eric Goldman in *Public Opinion Quarterly* (Goldman, 1944-45).

The Cantril survey provided a rich perspective on current opinion. In 1944 regularly conducted public national opinion polls were barely a decade old, and it was four years before their major preelection blunder. Nonetheless, many questions were quite appropriate to the issue of credibility in 1985, and here was an opportunity to see to what extent public opinion had changed over virtually the lifetime of the polling profession.

Considering opinions in 1944, a slim majority (56 percent) had even heard of public opinion polls and only 9 percent of the population reported that they followed them regularly. Forty years later, the percentage of the population that follows public opinion polls is much greater (Table 6).

Regarding the accuracy of polls, in 1944 a majority of those aware of polls felt that they were pretty nearly right most of the time in making

Table 7. The Accuracy of Preelection Polls

Some polling organizations make frequent predictions of election results. What is your general impression of how well they do—do you think they are pretty nearly right most of the time, or do you think their record is not very good?

	1944[a]	*1985*
	%	%
Right most of the time	57	$\boxed{68}$
Not very good	21	20
Don't know	22	12
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

Table 8. The Accuracy of Polls on Issues

Do you think poll returns on matters not dealing with elections, but with public opinion toward such things as labor problems or international affairs, are usually pretty nearly right or not right at all?

	1944[a]	*1985*
	%	%
Pretty nearly right	52	55
Not right at all	12	$\boxed{21}$
Don't know	36	24
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

election predictions. Opinion today is not very different. However, a larger percentage of those questioned in 1985 felt that polls are pretty nearly right most of the time (Table 7).

In 1944, compared to election predictions, it was less clear to the public whether the polls were accurate in dealing with issues. Note that over one in three had no opinion, but of those expressing a view, a *majority* felt the polls were pretty nearly right on such things as labor problems and international affairs (Table 8). Forty years later, the public is also less certain about the reliability of polls on issues, relative to election predictions, and among those expressing an opinion, the proportion saying "Not right at all" has increased significantly.

Finally, in 1944, the overwhelming majority of respondents judged polls as good for the country—but one in five had no opinion (Table 9). Forty years later, an equally large proportion felt polls are good things, but the percentage reporting they are bad things had increased somewhat.

Table 9. Overall Opinion Toward Polls

In general, would you say that polls of the opinions of the public are a good thing or a bad thing in our country?

	1944[a] %	*1985* %
Good thing	73	76
Bad thing	6	$\boxed{12}$
Don't know	21	12
	100	100
N responses	(1120)	(1023)

[a] Based on those aware of polls.

The results of these two surveys suggest that despite frequent criticism of the polls during the 1984 election, elite and public opinion of the polls was generally positive. There was certainly no evidence of a crisis in confidence. Criticism of the polls during the election either failed to reach large proportions of the public and press or was dissipated by the fact that the outcome of the election bore general resemblance to what most polls were showing. One can only speculate that had Reagan's margin of victory been significantly different from what most polls showed in their final days, the results of these surveys might have been very different.

From a historical viewpoint, opinion of the polls today among the public is not very different from what it was more than 40 years ago, when public polls were in their infancy. Unfortunately, there are no data points in between to demonstrate whether or how public opinion about the polls has varied over the past four decades.

The press survey did point to significant concerns among media respondents about the intrusiveness of polls in election campaigns, while the public survey did not cover this area. Supporters of Mondale had fewer positive things to say about the polls than did Reagan supporters.[2]

There are two other important qualifications to these positive findings for polls. Establishing the credibility of polls with a survey begs the question of whether nonrespondents to either survey had less favorable attitudes than those who participated. However, Gallup experience indicates that nonresponse to surveys is usually based on practical matters such as convenience, timing, interviewer approach, and so on. Opinions about polls, privacy, or other substantive criticism of surveys have tended to be relatively minor factors. It should also be noted that if these surveys were, in fact, subject to some bias because of selective participation, the bias would have had to be enormous to change the overall conclusions.

---

[2] For example, compared to Reagan critics, those who approved President Reagan were more likely to feel that polls are pretty nearly right most of the time (by a margin of 72 percent to 64 percent). A similar pattern was evident in each evaluative question.

A more troubling qualification to the press poll is the fact that the margins of favorability are wider than they are deep. As the tables above indicate, few respondents selected the most positive category in each of the interrogations about opinion polls. It is probable that the favorable attitudes expressed by press respondents could be easily changed, given the small proportion who expressed a high degree of confidence. It is also likely that had the poll been conducted prior to the election, results might have been quite different, given the softness of opinion.

To conclude, the surveys should provide a degree of comfort about the public perception of the reliability of opinion polling. On the heels of the 1984 election a positive impression existed; how well it holds up will depend in part on how well polling organizations perform in subsequent elections. The comments of the founders of the British and American Gallup Polls are quite relevant to this point:

Henry Durant:

> My advice to all would-be election forecasters is to think again. It is the most stupid job you can ever take up, no matter how hard you try to find a worse one. If you get the election right, everyone takes it for granted. If you get it wrong, you are standing alone and utterly ashamed, and there is nothing you can do about it.

George H. Gallup:

> It would be difficult to find any other profession whose leaders are held publicly accountable for their actions in quite the same way. However, it is the awful penalties for being "wrong" that have spurred the development of research procedures. As a result, in no other field of the social sciences has the prediction of human behavior reached the same high level of accuracy (Gallup, 1972).

References

Claffey, Charles E.
1984 "Pollsters did better, Gallup did best." Boston Globe, Nov. 11.
Gallup, George
1972 The Sophisticated Poll Watcher's Guide, 1972. Princeton, NJ: Opinion Press.
Goldman, Eric F.
1944–45 "Poll on the polls." Public Opinion Quarterly 8:461–67.

Haywood, Richard
1984 "Conflicting campaign polls point to one certainty: some are wrong." Washington Post, Oct. 31.
Rosenbaum, David E.
1984 "A good election for poll takers." The New York Times, Nov. 8.
Siegel, Barry
1984 "Polls: more art than science." Los Angeles Times, Aug. 27.

Evaluating Polls with Poll Data

BURNS W. ROPER

In an attempt to learn how the American public feels about opinion polls, The Roper Organization conducted a national poll in late March of 1985.[1] A total of 10 questions was asked. First people were asked if they had been interviewed before "this" interview, either by telephone or in person. Most had been, many of them several times (Table 1).

People were then asked whether they thought most opinion polls worked for or against the best interests of the general public (Table 2). Three out of four said polls worked for the public's interests—over nine times as many as said they worked against the public's interests.

[1] The survey was based on face-to-face interviews with a nationwide sample of 2000 adults age 18 and up. Down to the block level the sample was a probability sample of the 48 contiguous states. At the block level, sex, age, and employed women quotas were imposed, with evening and/or weekend interviewing required to insure proper representation of employed people. Since quota controls were imposed and specific respondents were not designated, "completion rate" is an inappropriate consideration in this sample.

The questions cited in this article were asked at about the 40-minute mark in a wide-ranging 45-minute interview—after questions about international affairs, social issues, lifestyles, attitudes toward business, public radio and television listening, and an assortment of other subjects. Since nearly half of all respondents had never been interviewed before, this subject matter—and length—had to affect their attitudes toward polls, particularly with respect to the questions reported in Tables 8, 9, and 10. A similar context and length would also influence responses to the question that asked about agreement with Letter X or Y: While more than half of all respondents had been interviewed before, nevertheless this interview was their most recent experience and undoubtedly had some effect on their responses.

Abstract The Roper Organization conducted two surveys in 1985 on the public's attitudes toward polls. One was a 10-question interview and the other was a single question asking for reactions to letters both critical and laudatory of polls. Results of the two surveys, which were asked of national samples of adults, indicate that although there is skepticism among the public about polls, few people are entirely negative. However, the rise in the last 10 years in the percentage of those who are critical of polls suggests that there is cause for concern. The author makes suggestions for improving polling from the respondent's point of view.

Burns W. Roper is Chairman of The Roper Organization, Inc.

Public Opinion Quarterly Vol. 50:10–16 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-10/$2.50

Table 1. Previous Interviews

Aside from this interview, about how many other times, if any, have you been interviewed in a poll—either in person or by telephone?

Never before	41%
Once	17
Twice	16
3 to 5 times	16
6 or more times	9
Don't know	1

Table 2. Do Opinion Polls Work for the Public's Interests?

Now I'd like to ask you some questions about polls in general, based on what you know or have read or heard about them; and not just your experience in this interview. First, do you think most opinion polls work for or against the best interests of the general public?

For	75%
Against	8
Don't know	18

Table 3. Influence of Polls

How much do you think poll results affect or change things—would you say polls usually have almost no influence, or some influence, or quite a lot of influence, or have too much influence?

Almost no influence	9%
Some influence	58
Quite a lot of influence	21
Too much influence	5
Don't know	7

Table 4. Accuracy of Polls

Do you think polls are almost always accurate, usually accurate, accurate only sometimes, or accurate very seldom?

Almost always accurate	10%
Usually accurate	46
Only sometimes accurate	32
Very seldom accurate	6
Don't know	6

How much influence do poll results have in terms of affecting or changing things? Most said "some" influence. Few said either none or too much (Table 3).

Next, how accurate are polls and how honest are pollsters? More said "accurate" than said "inaccurate," but not by an overwhelming margin (Table 4).

Table 5. Honesty of Pollsters

Do you think pollsters are almost always honest about their polls, are usually honest, honest only sometimes, or honest very seldom?

Almost always honest	25%
Usually honest	51
Only sometimes honest	14
Very seldom honest	3
Don't know	7

Table 6. Representativeness of Polls

Do you think a sample of 1500 or 2000 people can accurately reflect the views of the nation's population, or that it's not possible with so few people?

Can accurately reflect views of nation	28%
Not possible with so few people	56
Don't know	15

Table 7. Are Interviewees Typical?

Do you think pollsters interview typical, representative people, or that they get mostly unusual, nontypical, or even oddball types of people?

Typical, representative people	80%
Unusual, oddball types	7
Don't know	13

Pollsters fared better on their honesty than on the accuracy of their product. Many more said "honest" than said "not honest" (Table 5).

One of the most interesting questions in the survey was the next one, which asked people whether they think a sample of 1500 or 2000 people can accurately reflect the views of the nation's population, or that it's not possible with so few people (Table 6). By a two-to-one margin people said it's not possible. In fact, exactly the same percentage said that it is not possible as had earlier said that polls were either almost always or usually accurate (see Table 4)—56 percent in both cases.

Roper asked people whether they felt pollsters interview typical, representative people, or mostly unusual, nontypical people (Table 7). Eight in 10 said typical; less than one in 10 said atypical.

People were asked how many survey respondents they think tell the truth in polls (Table 8). Eight in 10 said most or nearly all people tell the truth. Very few think that very few tell the truth.

Table 8. Do Respondents Tell the Truth?

Do you think nearly all people who are interviewed in polls tell the truth, or most people do, or only some do, or that very few in a poll tell the truth?

Nearly all	28%
Most	53
Only some	13
Very few	3
Don't know	3

Table 9. Are Polls Enjoyable?

From the point of view of the person being interviewed, do you find polls an enjoyable and satisfactory experience, or an annoying and unsatisfactory experience, or somewhere in between?

Enjoyable and satisfactory	51%
Annoying and unsatisfactory	5
Somewhere in between	42
Don't know	1

To close the survey, respondents were asked whether they find polls an enjoyable and satisfactory experience, an annoying and unsatisfactory experience, or somewhere in-between. Half said enjoyable and satisfactory (Table 9). Nearly as many said somewhere in-between—which is fine, if in fact they weren't really like the disgruntled 5 percent, but too polite or timid to say so.

The final question asked the disgruntled 5 percent why they found polls annoying and unsatisfactory, and used a card that listed seven criticisms or complaints about polls (Table 10). The most common complaint, cited by nearly two-thirds of those who find polls annoying and unsatisfactory, was that they take too much of your time. (This question was asked at about minute #40 in a 45-minute face-to-face interview.) The second most common criticism, also related to time, was that polls interrupt you at inconvenient times. Fortunately, only 1 percent of the disgruntled said that the interviewers are often unpleasant. But, of course, telling that to the interviewer who was asking the question would require nerve, or bad manners, or both on the respondent's part.

This poll raises at least two questions. One is the matter of sample bias, and the other is the honesty of people's responses to these questions. Will those people who either dislike or disapprove of polls agree

Table 10. Why Are Polls Annoying?

Which of the reasons on this card, if any, are your reasons for finding polls annoying and unsatisfactory? Any others?

	% of the Total Sample	*% of Those Who Find Interviews Annoying and Unsatisfactory*
They take too much of your time	3%	63%
They interrupt you at inconvenient times	2	39
The questions are often too personal, nosey, and intimate	1	28
The questions are often stupid or unanswerable or both	1	26
The questions are often too difficult and complicated	1	24
Too often there is some other purpose to the poll, trying to sell you something or finding out when no one in your family is home	1	22
The interviewers are often unpleasant	—[a]	1
None	—[a]	2
Don't know	—[a]	1

[a] Less than 0.5%.

to be interviewed in the same proportions as those who are not critical of polls? Logic would say that more of the critics would fail to cooperate. And will the critics who agree to be interviewed tell it like it is, or will they pull their punches and shade their answers? How many of the 42 percent who said being interviewed is somewhere between enjoyable and annoying really meant annoying and unsatisfactory?

My judgment—guess, really—would be that sample bias and shaded responses could make the foregoing results 10 points more favorable than they really should be.

In the poll just described, respondents had eight opportunities to express one or more of 11 reservations about some aspect or other of polls. Three out of four respondents interviewed did, in fact, express one or another reservation, the single most common reservation being that you can't accurately reflect the nation's opinions with a sample of 1500 to 2000 people.

In addition to this poll, Roper has asked a single question on four different occasions in which respondents were shown excerpts of two letters that were actually received from survey respondents. One of these letters was highly critical of being interviewed and the other highly laudatory about the experience. At the end of four different surveys, we asked respondents to read the two excerpts and to tell us which letter they most agreed with. In the first three askings, which covered the mid- and late 1970s, 3 percent agreed with Letter X, the critical letter. More than 90 percent on each occasion agreed with the laudatory letter, Letter Y. However, on the latest asking, which was two years ago, agreement with Letter X had doubled to 6 percent; agreement with Letter Y had dropped from 94 percent to 86 percent, and the undecided rose to its highest level ever, 8 percent.

Excerpt from Letter X:

> On last Saturday, one of your representatives came to my home and asked if she could interview me.
>
> I now regret that I cooperated and granted the interview. I feel that it was an imposition to take that much time purely for the benefit of your clients. I also feel that some of the questions asked for my opinion about things that are entirely my own business.

Excerpt from Letter Y:

> I have just given an interview to one of your interviewers and wish to thank you for the opportunity to make my opinions known.
>
> I believe you perform a vitally needed service in times like these by giving people the chance to make their opinions known—and feel that eventually it will do some good. If leaders in industry and government can know what people really think they can do a better job.

There are the data. What conclusions can be drawn from them? The question as to what these data mean would be a matter of judgment, even if there were no question about what the numbers themselves really should be. And judgment, unlike sampling error, is not subject to a mathematical measure of error.

The two studies indicate—to me, at least—that we in the opinion research field don't get a clean bill of health, but we don't face a crisis of public confidence either. The studies say there is a healthy skepticism among the American public about polls—with equal emphasis on "healthy" and "skepticism." Most people are not automatically seduced, blindly taken in, and convinced by polls. On the other hand, few appear to be irate, up in arms, or angry over them.

We *do* face *some* problems; the suggested rise in agreement with Letter X, even if that increase is not statistically significant, is cause for concern.

There are areas in which we need to clean up our act a little. The biggest justifiable complaint that the public has with polls (that are in fact polls) is the matter of interview length. Face-to-face interviews of an hour or more and telephone interviews of a half hour or more are too long unless very special arrangements are made—somewhat too long, I believe, in terms of respondent fatigue and the validity of a respondent's answers, but substantially too long in terms of imposing on a respondent without his or her express consent.

Many researchers are sometimes guilty of asking questions that are too personal—or at least too personal for many subgroups of the population ("How often do you have a daily [sic] bowel movement?").

A third criticism that can be leveled at opinion researchers—and it's the number one criticism of phony calls—is taking advantage of respondents in terms of unexpected follow-ups or attempted sales. When a person answers that her refrigerator is 15 years old, that should not subject her to unwanted sales calls from an appliance dealer or manufacturer. When a person agrees to be interviewed, that does not mean that the survey organization has license to call back on three or four additional follow-up interviews over the coming year unless the respondent has agreed to such follow-ups.

There is not a serious problem now—fortunately—but as polling gets more and more pervasive, the problem will get worse, even if our actual practices don't change one whit. The same practices in more and more polls may be what is causing the agreement with Letter X to rise.

We have to monitor and regulate our practices. What would the impact be on our major excess if our code of ethics were changed? Suppose in the responsibilities-to-survey-respondents' part of the AAPOR code researchers were required to tell survey respondents within three minutes of the start of an interview how much longer that interview would take. Not "Could I ask you a few questions?" or "This survey will take just a few minutes." But "This survey will take a total of an hour and ten minutes. Is now a convenient time to talk to you, or would it be better if I call back at another time?" If we were required to so inform survey respondents, might not interview lengths drop appreciably? Even if they did not drop appreciably, survey respondents would probably feel better about having been told up front. They would feel less like they had been had.

Trends in Attitudes Toward and Participation in Survey Research

STEPHEN SCHLEIFER

Because the survey research industry depends on the public's willingness to participate in its surveys, public goodwill must be a top priority for the nation's survey research community. To this end, Walker Research Inc.[1] has since 1974 conducted a biennial Industry Image Study, which examines the public's attitudes toward the survey research industry and some of its specific practices. Interviews are conducted with a national sample of respondents to determine their frequency of participation in polls and surveys, their opinion of the interview experience, as well as their feelings about the research industry in general. These biennial Industry Image studies have monitored trends in public opinion of survey research.

This article summarizes the findings of the 1984 Industry Image Study, which measures participation levels in surveys, attitudes toward the interview experience, and feelings about the survey research industry. (See Appendix for question wording.) The results and the trends observed from prior years provide clues on how the industry can enhance its image among the general public.

METHOD

In the week of August 20–26, 1984, 495 telephone interviews were completed[2] with a national probability sample of all U.S. telephone-

---

[1] Walker Research, Inc., is the fourteenth largest marketing research firm in the country, according to *Advertising Age* (5/23/85). Headquartered in Indianapolis, IN, Walker has eight branches nationwide and employs over 1000 people.

[2] Interviews were conducted from Walker Research's National Calling Center located in Indianapolis, IN, using its automated CRT Interviewing System. Three attempts were made on each sample household using a day-night calling pattern. The national sample was randomly generated by computer to include both listed and unlisted telephone numbers.

Abstract This study, now in its sixth wave, examines trends in consumers' willingness to be surveyed as well as attitudes toward being interviewed and toward the survey research industry in general. The results raise some important questions in terms of multiple participation and its effect on the validity of survey data. They also challenge researchers to examine their own standards and practices in order to enhance the relationship between researchers and the participants.

Stephen Schleifer is Senior Vice President of Walker Research, Inc. of Indianapolis, IN. This article is a revised version of a presentation given at the 40th Annual Conference of AAPOR at McAfee, NJ, May 18, 1985.

Public Opinion Quarterly Vol. 50:17–26 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-17/$2.50

Table 1. Limitation of Findings/Refusal Summary[a]

Representative Sample	*1978*	*1980*	*1982*	*1984*
Completed interviews (*N*)	500	500	499	495
% Complete	51%	59%	54%	54%
Total refused	477	346	432	415
% Refused	49%	41%	46%	46%
At introduction	22%	16%	30%	40%
During interview	27%	25%	16%	6%

[a] The findings of the 1984 Industry Image Study are limited by the normal statistical error inherent in any sampling process. Because the research was conducted by telephone, it is not necessarily representative of nontelephone-owning households. Further, increased potential for nonresponse bias exists in the findings due to the relatively high level of respondent refusals (as has been observed in past studies).

owning households. Of these, 115 people (23 percent) reported participation in at least one survey in the past 12 months. To increase the number of these past year participants, additional interviews were completed. Therefore, the "representative sample" includes 495 respondents, including 115 who claimed survey participation in the past year. The "past year participant sample" includes the 115 respondents chosen at random, *plus* 86 obtained on an over-quota basis (for a total of 201 people).

The first two waves of the Industry Image Study (in 1974 and 1976) included interviews from only the top 30 markets. In 1978, the sampling plan was modified to a national probability design. The 1984 results can therefore be directly compared to the 1978, 1980, and 1982 surveys. The response rates achieved since 1978 are shown in Table 1.

RESULTS

Survey Participation. Annual studies conducted by industry groups such as the Council of American Survey Research Organizations (CASRO) and the American Marketing Association (AMA) estimate participation levels in surveys based on figures provided by research organizations and manufacturers. By contrast, this study's projections emanate from the reports of survey participants themselves. According to the 1984 results, overall participation in polls and surveys has remained constant since 1982, which was down somewhat from 1980. As illustrated in Figure 1, 54 percent of the representative sample said they had been interviewed at some point but not in the past year, while 23 percent said they had been interviewed within the last 12 months. Identical results were obtained in 1982. By contrast, 1978 levels were

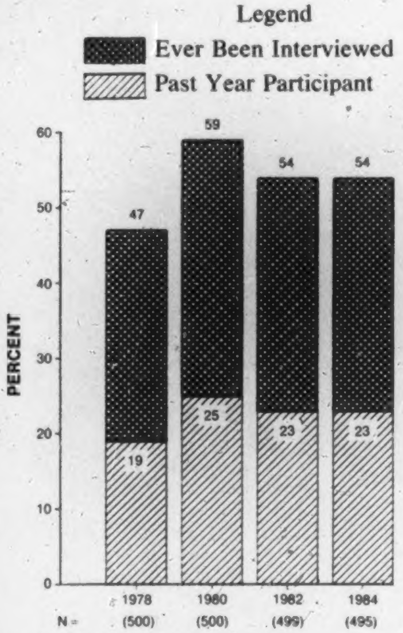


Figure 1. Participation Levels Among the Representative Sample

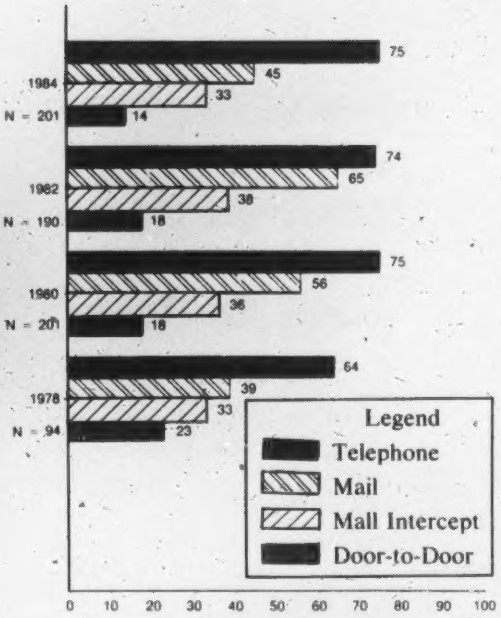


Figure 2. Participation by Type of Interview: Past Year Participation

significantly lower[3] than in any of the following years. These findings coincide with the data published annually in *Advertising Age* showing the real growth of survey research slowing from the explosive expansion observed in the late 1970s.

The type of interviews respondents participated in over the past year remains basically unchanged from prior measurements, as shown in Figure 2. According to 1984 past year participants, telephone continues as the dominant form of data collection (75 percent) followed by mail (45 percent), mall intercept (33 percent), and door-to-door (14 percent) surveys. After experiencing rather dramatic increases in recent years, mail surveys showed a significant *decrease* in 1984. Door-to-door interviewing also continues to decline as a viable method.

With this significant amount of survey activity, concern has been voiced about the "freshness" of those sampled. Are researchers reaching new participants each time, or are they interviewing the same respondents survey after survey? Prior years' results suggested that multiple participation is occurring and might be growing. To quantify this hypothesis, 1984 respondents who said they were interviewed in the

---

[3] Throughout this article, the term *significant difference* reflects a significant change using an independent *t*-test at the 95 percent confidence level.

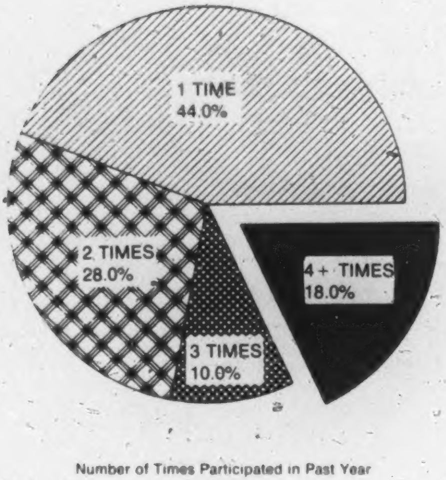


Figure 3. Repeat Participation in Past Year

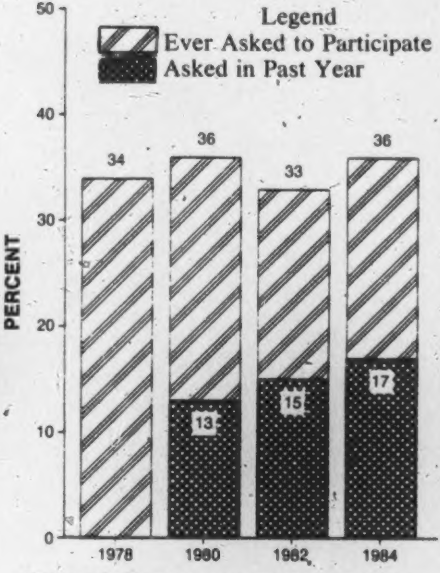


Figure 4. Exposure to a Disguised Sales Pitch

Table 2. Multiple Participation in Single Method

	1978	*1980*	*1982*	*1984*
Base: past year participants (*N*)	94	201	198	201
	%	%	%	%
Telephone	51	56	43	50
Door-to-door	22	42	30	43
Mall intercept	25	43	48	26
Mail	43	57	58	47

past year were asked how often. Figure 3 shows the majority of past year participants (56 percent) said two or more times, while 44 percent said they were surveyed just once over the past 12 months. Of concern is that 18 percent of the past year participants said they had been interviewed four or more times.

Further questioning into repeat participation by method type revealed a reversal in 1984 from the trends observed in 1982, as shown in Table 2. After a significant decrease in 1982 (from 1980), repeat participation in telephone interviews rebounded somewhat in 1984, returning to the level first observed in 1978. As for mail surveys, this research reveals a significant decrease in multiple participation comparing 1984 data to 1980 and 1982. Another decrease over the past year is seen for mall intercept studies. After a steady rise in reported repeat participa-

Table 3. Demographic Profile of Past Year Participants

	1984 Participants	*1984 Nonparticipants*
Base: Total respondents (*N*)	201	380
Education	%	%
High school grad or less	46	60
Some college or more	52	36
Income		
Under $20,000	26	36
$20,000 plus	51	41
Age		
Under 35	43	27
35 plus	54	67
Sex		
Male	21	46
Female	79	54

tion, the significant decrease found in 1984 will be encouraging to some, and may perhaps reflect an increased emphasis to screen out frequently intercepted respondents.

It is difficult to draw definitive conclusions from these data on repeat participation. However, they suggest the need to address some important questions:

> Are there segments of consumers more willing to cooperate, who consequently tend to be overrepresented in surveys?
>
> How much is "too much?" At what level of repeat participation should researchers become concerned?
>
> To what extent does more frequent participation in polls and surveys create "professional" respondents rather than typical consumers? What is the effect of multiple participation on the validity of the data collected?
>
> Should previous participation become a standard screening question for polls and surveys? Is it worth the expense to include only "fresh" respondents each time?

These and other questions should be addressed by those of us engaged in survey research. While we can only speculate about the characteristics of people who steadfastly refuse to participate in surveys, the Industry Image Study does reveal some things about those who do. Table 3 compares the demographic characteristics of past year participants with those who had not been surveyed over the same period of time. Past year participants in 1984 were significantly more likely than nonparticipants to be under 35 (43 percent versus 27 percent), to have completed at least some college (53 percent versus 36 percent), to have income over $20,000 (51 percent versus 41 percent), and to be female (79 percent versus 54 percent). These demographic differences are consistent with those reported in previous waves of this research.

The increasingly targeted nature of marketing researchers' samples and the entire industry's continued growth may in part explain the high levels of repeat participation in surveys. While public opinion surveys typically sample a cross-section of the general public, marketers' growing use of market segmentation requires that they look for specific population groups which are thought to be the best consumers for their goods or services. This study suggests that some of these groups—the young, better educated, higher income, females—are of particular interest to many, and that they are oversurveyed. If the trend continues, it may become more difficult to interview these popular segments, and both refusal rates and repeat participation may rise.

False Surveys. A factor contributing to the public's attitudes toward the research industry is false surveys—selling efforts under the guise of survey research. According to the 1984 Industry Image Study, reports of such false surveys remain at relatively high levels. As shown in Figure 4, over one-third (36 percent) of those in the representative sample said they had at some time been asked to participate in an alleged poll or survey which turned out to be a sales pitch. The trend is moving upward to the point where 17 percent of 1984 past year participants said they were subject to one or more false surveys. These disguised sales pitches harm survey researchers' reputations, raise the public's suspicions, and contribute to increased refusal rates.

Attitudes Toward Being Interviewed. Based on most industry projections, the level of survey activity will continue to grow as the demand for feedback from the marketplace accelerates. Researchers' ability to meet this demand depends on the public's attitudes about being surveyed. What, therefore, does the public think of survey research which asks much of them, and may appear to give little or nothing in return?

The 1984 results show a downward trend in those past year participants (79 percent) who rated their last interview as a "pleasant" experience (see Figure 5). In addition, increasing numbers said that the interview was too long (23 percent) and contained what participants felt were confusing (16 percent) and overly personal (14 percent) questions. As shown in Figure 5, the percentage of past year participants who said their last interview was an "unpleasant" experience rose significantly from 6 percent in 1982 to 12 percent in 1984. While the other shifts from 1982 to 1984 in Figure 5 are not statistically significant, the directional implications are clear and consistent. The data do suggest, however, that researchers can exercise some control

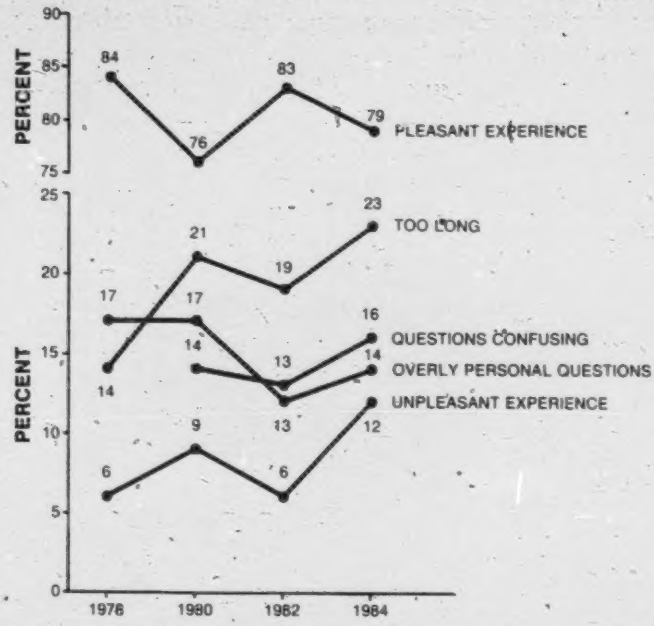


Figure 5. Attitudes Toward Being Interviewed: Last Interview Evaluation

over respondents' attitudes. Shorter interviews, clearly worded questions, and the elimination of unnecessary personal or sensitive items would help. Determining what is an optimum interview length or which questions are perceived as too personal are issues for further study.

There is reason to believe, however, that the element most critical in shaping respondents' attitudes is the interviewer. For the first time in 1984, past year participants were asked why they had evaluated their last interview as either "pleasant" or "unpleasant." The results showed that the personality or attitude of the interviewer was cited most frequently (47 percent) by those who said it was a pleasant experience. Other potential influencing factors such as questionnaire design (20 percent), the study's particular topic (8 percent), or whether the interview was conducted at a convenient time (3 percent) were mentioned less frequently. Only 1 percent (3 respondents) said the experience was unpleasant. These findings have obvious implications for interviewer and supervisor training, practices, and procedures.

Public Attitudes Toward the Industry. The public's overall attitudes toward the survey research industry are somewhat more positive than those expressed about their actual interview experience. These attitudes were assessed by asking respondents to agree or disagree with a series of belief statements. As shown in Table 4, the majority (81 percent) of respondents feel the research industry serves a useful purpose,

Table 4. Image of the Industry

	1978	*1980*	*1982*	*1984*
Base: representative sample (*N*)	500	500	499	495
Positive image attributes	%	%	%	%
The research industry serves a useful purpose	85	85	84	81
Polls and research surveys are used to help manufacturers produce better products	81	80	80	82
Answering questions in polls or surveys is an interesting experience	70	64	66	60
Answering questions in polls or research surveys is in my own best interest	72	70	68	60
Survey research firms maintain the confidentiality of answers	—	60	61	54
Negative image attributes				
Some questions asked in polls or research surveys are too personal	42	47	49	51
The information obtained in polls or research surveys helps manufacturers sell consumers products they don't want or need	38	45	43	44
The term *poll or research survey* is used to disguise a sales pitch	43	38	38	39
Polls cr research surveys are an invasion of privacy	31	25	26	28
Answering questions in polls or research surveys is a waste of time	17	18	21	22
The true purpose of some surveys is not disclosed	—	68	65	61

and that surveys are used to help manufacturers develop better products (82 percent). On the other hand, only 54 percent of the 1984 representative sample believes survey research firms maintain the confidentiality of answers, compared to 61 percent who held this belief in 1982. This finding is particularly disturbing because researchers must have the public's trust to gain access and obtain valid data.

Overall, based on the answers to these industry-related questions, the signals being sent to researchers are mixed. On the one hand, some attitudes suggest the public doesn't object to surveys and polls *in principle*. On the other hand, many people remain suspicious of the industry's methods and motives.

CONCLUSIONS

The data obtained in this ongoing study of attitudes towards the survey research industry are generally encouraging, although several problem areas do exist. These problems must be addressed by *all* segments of the research industry in order to maintain and improve the quality of data collected via polls and surveys. Because the industry depends on public cooperation for its livelihood, efforts must continue to solidify the relationship between researchers and potential respondents. The results of this survey indicate that while the public is supportive of the research industry in general, some of its perceived practices appear to be eroding these positive feelings. Interviewers who are poorly trained and supervised, interviews that are too long and contain overly personal questions, and repeated requests to participate in surveys appear to be among the causes. To reverse this, all those engaged in survey research must continually evaluate their own practices from the *respondent's* point of view, and carry out their investigations in a way that will not harm the industry's image.

APPENDIX: 1984 INDUSTRY IMAGE STUDY

1. Thinking back, have you, yourself, ever participated in a survey before, or not?

2. Have you, yourself, been interviewed for a survey in the past year, or not?

3a. Thinking about the past year, how many times have you, yourself, been interviewed for a survey?

3b. And now, thinking *just of the past year*, about how many times have you, yourself, been interviewed for a survey. . . By telephone? In person at home? In person at a store or shopping center? By a questionnaire sent to you in the mail?

4a. Now I'd like you to think of the *last* time that you participated in a poll or research survey. Was that particular interview by telephone, in person at home, in person at a store or shopping center, by a questionnaire sent in the mail, or by some other method?

4b. Still thinking of the *last* time you were interviewed, were any of the questions worded in a confusing manner, or not?

4c. Were you asked any questions that you thought were overly personal during that last survey, or not?

4d. What did you find overly personal about those questions?

4e. About how long, in minutes, did that last survey take?

4f. In your opinion, did the survey last too long, or not?

4g. And, still thinking of the last time you were interviewed, was your *overall* impression of the interview pleasant or unpleasant?

4h. You said your overall impression of the interview was *pleasant*. Why did you say that? What other reasons? Why else?

4i. You said that your overall impression of the interview was unpleasant. Why did you say that? What other reasons? Why else?

4j. What was the topic or subject of that last interview?

5. Now, I'm going to read a list of statements which some of the people we've talked with have said about polls or research surveys. As I read each statement, please tell me whether you agree or disagree with the statement.

STATEMENT

The survey research industry serves a useful purpose.

Answering questions in polls or research surveys is in my own best interest.

Answering questions in polls or research surveys is a waste of time.

The term "poll or research survey" is used to disguise a sales pitch.

The information obtained in polls or research surveys helps manufacturers sell consumers products they don't need or want.

Some questions asked in polls or research surveys are too personal.

Polls or research surveys are used to help manufacturers produce better products for consumers.

Polls or research surveys are an invasion of privacy.

Answering questions in polls or research surveys is an interesting experience.

Poll or research survey companies should not be allowed to interview people in shopping centers.

Survey research firms maintain the confidentiality of people's answers.

Research surveys give people an opportunity to provide feedback to manufacturers and retailers about the products or services they provide.

The true purpose of some surveys is not disclosed to the people who participate.

6a. In the past year, have you refused to participate in what appeared to you to be a legitimate poll or research survey, or not?

6b. Have you *ever* been asked to answer questions in a poll or research survey that turned out to be a sales pitch for a product or a service?

6c. Have you been asked to answer questions in a poll or research survey which turned out to be a sales pitch for a product or service *in the past year*, or not?

6d. What types of products or services did the sales pitch deal with?

7a. Surveys are conducted on a wide variety of subjects or topics. I would like to read a short list of different topics and ask you how likely you would be to participate in such a survey.

The first/next topic is (*TOPIC*). How likely would you be to participate in a survey that asked (*STUDY*)? Would you say you would participate, might or might not participate, or would not participate?

TOPIC (STUDY)

Automobiles (for your reaction to new automobile designs).

Political candidates (who you plan to vote for in the next election).

Food products (for you to try a food product and get your opinion of it).

Advertising (for you to recall what advertisements you have seen recently and what you think of them.

Television (you to describe your normal television viewing habits).

Telephone company (you to evaluate the service you get from the telephone company).

Personal hygiene products (for you to evaluate a specific brand of personal hygiene product).

7b. The telephone is used for a variety of purposes, including surveys as well as for other purposes. I would like to read a short list of the ways the telephone is sometimes used and ask you how acceptable each would be to you if you were called for such a purpose.

One/Another way the telephone is used is (*PURPOSE*). How acceptable would such a call be to you? Would you say it would be very acceptable, somewhat acceptable, or not at all acceptable?

PURPOSE

Calls from the blood bank asking you to donate blood.

Charitable organizations asking for contributions.

The selling of insurance.

Handicapped people selling household products.

Organizations such as Goodwill or AmVets asking you to donate used goods.

Real estate agents looking for business.

Companies selling home improvement products such as aluminum siding.

Political organizations asking you to vote for specific candidates.

Surveys or research companies asking you for your opinion on a product or service.

Now I have just a few last questions to help us group your answers with those of others who might be somewhat like you.

[A–D are demographic.]

E. Finally, how long would you estimate that this interview lasted?

F. Do you consider that to be "too long" or not?

Surveys on Surveys: Limitations and Potentialities

JOHN GOYDER

A "survey on surveys" questions respondents about attitudes and past actions concerning surveys, the pioneering exemplar being Sjoberg (1955).[1] Several aspects of such data have interested survey researchers: The extent of bogus surveying has generated sustained attention (Arnold, 1964; Baxter, 1964; Biel, 1967; Allen and Colfax, 1968; Rugg, 1972), and questions about preferred form of contact appeared in investigations of the acceptance of telephone surveying (Rogers, 1976; Groves, 1979:193). Attention to sensitive question topics was reported in Singer (1984), measures of "response burden" have been fielded by Sharp and Frankel (1983), and an emphasis on recording motives for refusal appears in studies such as DeMaio (1980). Day (1975:464) discussed the legitimacy the public attributes to surveying. Schwartz (1964) collected data on most of the above themes, but perhaps the most comprehensive survey on surveys is a U.S. Census Bureau data set reported in National Research Council (1979).

---

[1] I have adapted the term from Sjoberg's (1955) "Questionnaire on Questionnaires."

Abstract Data from two surveys on surveys in Waterloo, Ontario were used to assess attitudes toward the survey. Questions in the two surveys inquired into motives for refusing/consenting to interviews; attitudes toward past experience with surveys; likes/dislikes of form of contact, question format, and topic; perception of social pressure to respond to surveys; views on the social legitimacy of polling and on legislative controls for surveying. The surveys also collected response histories. Analysis of the attitudes began with a factor analysis of 12 items, which identified four factors descending from the general to the particular. It was found that attitudes are related to exposure to surveying, i.e., the more times a person reports requests for survey cooperation, the more unfavorable is his or her attitude toward the method. On the whole, however, the author finds that the evidence presented belies pessimistic views about surveys on surveys.

John Goyder is Associate Professor, Department of Sociology, University of Waterloo, Ontario, Canada. This article is a revision of a paper presented at the 40th Annual AAPOR Conference, May 16–19, 1985. Research assistants on the record-linking portion of the study were Joti Sekhon, Christine Hutchinson, Ella Haley, and Joan Lyons. Twenty-one students enrolled in Sociology 282 in Winter 1982 collected the 1982 survey data. Patricia Anderson, Helen Chapman, and Joan Lyons conducted telephone interviews for the 1985 replication. Marnie Goyder was coder on the replication. The work reported herein has been supported by the University of Waterloo Faculty of Arts Resource Fund, the UW-SSHRC Small Grants Subcommittee, and by an SSHRC Leave Fellowship. The hospitality of the Social and Political Sciences Committee, Cambridge University, in granting the author Visiting Scholar privileges in 1985–86, greatly assisted the completion of the manuscript.

Public Opinion Quarterly Vol. 50:27–41 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-27/$2.50

A new wave of survey on surveys data, reported in this issue of *POQ*, makes a reassessment of the limitations and potentialities of the data timely. By some criteria, the pool of survey data on surveying has become impressive. Where earlier research was often conducted on a local scale, and on a small budget (e.g., Sjoberg, 1955), many of the recent data derive from premier market research firms such as Gallup and Roper. The new studies thus often read public opinion on surveying at the national level, with replication at least of main themes, if not of specific questions, across survey houses.

On the debit side, however, the mounting quality of survey on surveys data serves only to sharpen more fundamental reservations about such investigations. Receiving much attention in the recent reports is the problem of "professional respondents" who, with nonrespondents by definition absent from tabulations, bias results in favor of public approval of surveying.[2] Another concern, potentially the more damaging, is the likelihood that public attitudes toward surveying amount to no more than unstable surface opinions, rather than definable attitudes in any psychological sense (Rokeach, 1968:455).[3]

The present article reports material that addresses these two aspects of the validity of surveys on surveys. The data are confined to one municipality but, as we shall argue in the following description of research design, they complement the national studies by addressing validity threats explicitly.

DESIGN

The epistemological limitation to surveys on surveys is self-evident; employing an instrument to measure its own performance is immediately contradictory. To assess attitudes toward the survey via attitude surveys is not unlike seeking to comprehend the mechanics of the camera solely from photographs. The contradiction becomes at least partially resolved, however, when survey data are placed alongside external sources of evidence, and in the present case this design was achieved by record linking.

Two surveys on surveys were collected in Waterloo, Ontario, Canada. The first, fielded in March 1982, employed student interviewers who attempted contact at the door with 140 names randomly (cluster) sampled from the city directory.[4] In April 1985 the survey was replicated, with professional telephone interviewers attempting contact with another 140 listings, drawn by the same method as in 1982. For all sample listings, respondents and nonrespondents alike, a data file was compiled and record-linked with the survey data. From municipal taxation records it was possible to learn property assessment, age, citizenship, and religion (Protestant–Catholic). The city directory added marital status, labor force standing, gender, some details on household composition, type of dwelling, ownership or rental of dwelling, occupation, migration history (from back issues), and a code identifying householders who refused to be enumerated for the directory. Since only a percent or two of the population refuse the directory, and because such refusals offered valuable leverage to the design, the 1985 sample was stratified in the proportion one-third city directory refusers to two-thirds random cross-section.

---

[2] "Melvin D. Field, the head of Field Research of San Francisco, said he felt that the studies were flawed because those people who answer questions in a poll are inherently more sympathetic to polls than those who refuse to be interviewed." (*New York Times*, 1985:B7).

[3] "Our image may only be as good as our last election." Andrew Kohut, President of Gallup, quoted in *The New York Times* (1985:B7).

Question design in the 1985 survey was identical to the 1982 schedule, excepting the deletion of a few of the less useful 1982 questions in the interests of streamlining the telephone enumeration. In final (1985) form, the interview schedule required some 15 minutes to administer and inquired into: attitudes toward past experiences with surveys; motive for refusing and consenting to interviews; likes and dislikes concerning form of contact, question format, and question topic; perception of social pressure to respond to surveys; general views on the social legitimacy of opinion polling; and attitudes on legislative controls on surveying. The attitude questions were most directly inspired by Sjoberg (1955), from whom several wordings were reproduced, but most of the topics investigated have appeared in the other literature cited above.

In addition to the attitude questions, the surveys collected response histories, reconstructing past contacts by surveyors, and eliciting details on the ensuing response decisions. We regard the response history as partially external to the attitude questions. Although collected within the same instrument as the attitudes, and thus subject to the same nonresponse bias and other errors, the response history constitutes a behavioral variable against which the attitude items can be validated.

---

[4] Time and resources permitted only two systematic calls, with the occasional third call to fulfill an appointment. For the 50 cases in which contact was not achieved, the survey instrument was mailed out. It had been designed to be self-administerable, since it was known that multiple callbacks at the door would be impossible. Fuller fieldwork design particulars for the 1982 survey appear in Goyder (1985a).

The overall case base for the analysis is small (280 sample listings, totaled over the two surveys, with 172 responses to the surveys), making .05 significance hard to achieve. Where key computations are at stake below, successful reproduction across samples of the direction of a relationship is viewed as supporting evidence given a marginal ($p < .1 > .05$) result on the formal significance test.

The article begins with the problem of "professional respondents," exploiting the record-link file for a nonresponse analysis of the Waterloo surveys. An assessment of the validity of attitudinal data about surveys is then attempted, using two criteria. The stability of factor analysis results for the attitude items from the two Waterloo surveys is tested by comparing loadings from the two data sets, and then the attitude constructs from factor analysis are validated against the history of response behavior.

RESPONSE BIAS IN THE SURVEYS ON SURVEYS

Fifteen of the total of 280 sample addresses drawn for the 1982 and 1985 surveys proved vacant and are removed from further analysis. Response on the remaining case base of nonvacant households was 65 percent in each sample. The telephone contacts involved up to five calls, a plan consistent with available resources and deemed likely to yield a response in the sixties. It seemed consistent with accurate replication to engineer equivalent response on each sample. Tabulations for analysis comparing respondents against nonrespondents appear in Table 1.

Logistic regression is employed, since response is binary. Thus, the dependent variable is the natural log of $p/(1 - p)$, where p is probability of response, and the regression coefficients entered into Table 1 are computed on the maximum likelihood criterion (Hanushek and Jackson, 1977:200–3).

Entered individually, as zero-order effects, seven sample listing characteristics passed .05 or .01 significance. (Given the small case base, we adopt in Tables 1 and 4 the convention of describing $p < .05$ as "significant," $p < .1 > .05$ as "marginally significant," $p < .5 > .1$ as nonsignificant but worth entering in the tables, as possibly reproducible in a larger data set, and $p > .5$ as so certainly due to chance as to not justify reporting). Since zero-order effects among sociodemographic characteristics are likely to be reexpressions of one another, a multivariate model was next computed, in which the independent variables were successively entered alongside occupational SES. SES is operationalized through two indicators: the Blishen scale (Blishen and McRoberts, 1975), a Canadian equivalent of the Duncan scale, and a dummy variable identifying cases where occupation was either uncod-

Table 1. Nonresponse Analysis, Waterloo Surveys

	Logistic Regression Coefficients (Dependent Variable = Proportion Responding)		
		Multivariate	
Predictors	*Zero-order*	*Net of SES*	*Net of SES and Noncontact*
Intercept	—[a]	—[a]	—[a]
SES	.039***	—[a]	—[a]
No data, SES	−.527**	—[a]	—[a]
Age	−.019**	−.018*	−.027***
City directory refuser	−.339		
Religion	−.261*	−.275*	−.285*
Marital status	.854***	.539*	.358
Labor force status	.343		
Head of household	−.429		
Family type	.689		
Gender	†		
Own or rent	.532**	.268	†
Property value	.0002***	.0001	.0001
Citizenship	−1.117		
Residential stability	†		
Apartment dweller	−.839***	−.663**	−.296

[a] Not entered—differs for each computation.

* $p < .1 > .05$

** $p < .05 > .01$

*** $p < .01$

No star, $p < .5 > .1$

† Not entered, since $p > .5$

Scoring notes: SES, No data, SES: see text/Age: years/City Directory Refuser: 0 = no refusal, 1 = refusal/Religion: 0 = Protestant, 1 = Catholic, but Protestant school supporter, 3 = Catholic, Catholic school supporter/Marital Status: 0 = single (incl. divorced; widowed), 1 = married, spouse present/Labor Force: 0 = not in labor force, 1 = in labor force, incl. students/Headship: 0 = not head, 1 = head/Family Type: 0 = one generation listed, 1 = multiple generation/Gender: 0 = female, 1 = male/Own or Rent: 0 = renter, 1 = owner/Property Value, in dollars/Citizenship: 0 = alien, 1 = citizen/Stability: 0 = different listing in city directory two years prior to survey, 1 = same listing/Apartment: 0 = individual entry dwelling, 1 = apartment with common entry.

N = 265, with means inserted for missing data.

able or unknown.[5] The logic behind preentering occupational SES is that socioeconomic status constitutes the most enduring of all themes in studies of survey response bias (e.g., references in Goudy, 1977). No causal assumption is intended; the interest in the present analysis simply lay in identifying response biases net of the ubiquitous "middle class" bias encountered in virtually every survey. Net of SES, response remains significantly (.05) lower among apartment dwellers and marginally (.1) lower among Roman Catholics, the unmarried, and the elderly.

---

[5] Thirty-three percent of cases lacked an SES code, the largeness of the figure being partly due to the oversampling of city directory refusers. Entries in city directories are hard to code onto a fine-grained scale, since the reports often simply identify a firm. That problem was attacked by checking back entries of the directory. Often, the entries from several years summed to sufficient information for an accurate code. A bias accumulates from the back-checking because marginals to the local labor force, as well as recent immigrants, become underrepresented. The dummy category was created to offset that bias.

In the final computation in Table 1, a dummy variable identifying confirmed contact was included in the model. The dummy virtually redefines response rate into refusal rate, since that part of response attributable to noncontact is partialed away. Apartment dwelling now drops from the model, confirming the intuition that low response from apartment buildings is attributable to the difficulty surveyors encounter in gaining access. Marital status, marginally significant in predicting total response, also slips below the $p = .1$ level. Again, the conclusion can be that response among marrieds is high for accessibility reasons (spouses suggest suitable moments to call back, etc.). Old age, however, gains strength as a correlate of low response and now surpasses .05 significance. One concludes that surveyors gain access to the young with difficulty, and this problem of access masks a relatively low refusal rate. Refusal on the city directory, the variable stressed earlier as tapping externally defined survey resistors, does not emerge as an important factor in the model of nonresponse. The effect takes the expected direction of city directory refusers tending to decline the Waterloo surveys, but falls well short of even .1 significance. The directory refusers receive further analysis below, but in the next section attention moves to the factor structure of attitudes toward surveys.

ATTITUDES ABOUT SURVEYS, AND SURVEY RESPONSE BEHAVIOR

Six questions from the Waterloo surveys were germane to constructing a general attitudes-toward-surveys variable: question 9, asking about objections to being asked questions about various topics; question 15, entailing eight agree-disagree rankings of statements about opinion polling, over a five-category scale; question 16, regarding city control of surveying; question 17, asking "should governments be the only ones allowed to conduct surveys;" question 19, inquiring whether the census should be voluntary or compulsory; and finally question 20, a Likert scale ascertaining favorability toward "the City of Waterloo conducting a regular poll on issues facing the city."

Details on question wording, together with descriptive statistics, appear in Table 2. A fuller descriptive account of the reading of public opinion toward surveying appears elsewhere (Goyder, 1985b), and for present purposes we shall simply note that, like other such informants

Table 2. Factor Analysis Results, for Respondents on Waterloo Surveys

Item	*Range*	*Mean*	*s.d.*	*High Score =*	*Factor Loadings*			
					1	*2*	*3*	*4*
For each of the following, could you say if you would object strongly, a little, or not at all, about being asked questions in a survey. (Score is total over 7 items, described in text.)	7–21	16.0	3.4	No objection	.152	.176	−.055	.279
Could you tell me whether you would strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree, with the following statements about political opinion polls:								
Public opinion polls make the country more democratic.	1–5	3.2	1.2	Disagree	.163	.007	.486	−.197
The biggest problem with opinion polls is the way the media distort the meaning of results.	1–5	2.1	1.1	Disagree	−.044	.269	.163	.113
Most opinion surveyors have faked their results on one occasion or another.	1–5	3.0	1.1	Disagree	.076	.730	−.114	−.007
Government pays too much attention to opinion polls.	1–5	2.8	1.2	Disagree	.110	.058	−.110	.535
Opinion surveyors interview too few people for results to be accurate for the whole population.	1–5	2.4	1.2	Disagree	.130	.364	.046	.119
Opinion surveyors are just out to make money.	1–5	3.7	1.0	Disagree	.480	.204	−.077	.128
Opinion polls are an invasion of privacy.	1–5	3.5	1.2	Disagree	.601	.075	−.105	.047
Opinion polls do no harm.	1–5	2.8	1.2	Disagree	−.328	−.189	.499	.068
Do you think the city should set up controls to restrict the interviewing of people?								
Should governments be the only ones allowed to conduct surveys? (Score is summation of the 2 questions.)	2–4	3.8	0.5	No	.419	−.051	.160	.063
Some people would say that the census, which Statistics Canada collects, should be voluntary but others feel it should remain compulsory, as it now is. Would you say the census should be voluntary or compulsory?	1–2	1.8	0.4	Compulsory	−.022	.054	.213	.213
Would you be very much in favor, somewhat in favor, neutral, somewhat not in favor, or very much not in favor of the City of Waterloo conducting a regular poll on issues facing the city?	1–5	2.2	1.2	Not in favor	−.009	.063	.390	−.038

(Kohut, 1986; Roper, 1986), the Waterloo respondents return a mixed verdict on surveying. Means tend to be centered in the middle of distributions, and standard deviations are relatively high. Only on the two questions about government control of surveying is there near-unanimity: the majority of respondents in both samples did not think the city should control interviewing, and disagreed that only governments should be allowed to conduct surveys. Responses for these two questions, both skewed binary variables, were summed to give a statistically better behaved index of government regulation sentiment. Topics (money, politics, religion, family life, occupation, sex, environment) in the battery to do with objectionable question domains also were summed, to provide a generalized index of level of objection to being asked survey questions.

The righthand portion of Table 2 describes factor analysis loadings for the 12 attitude items. The computational method is principal axis factoring, and varimax rotation, using SPSS-X. Missing data seldom exceed 1 percent in the Waterloo surveys, and were handled by substituting means.

The analysis identified four factors and, in characteristic pattern, the factors tend to descend from the general to the particular. Thus, factor 1 describes those who agree that surveys invade privacy, and that "opinion surveyors are just out to make money," while disagreeing with the statement "opinion polls do no harm." Factor 1 also incorporates the feeling that government should control surveying (all loadings .3 or greater). Factor 2 exhibits high (defined as .3) loadings on just two indicators: agreement that "most opinion surveyors have faked their results on one occasion or another" and that "opinion surveyors interview too few people for results to be accurate for the whole population." The two items suggest a "methodological factor" describing disbelief in the accuracy of polls, rather than dislike of surveying itself. Polling as legitimate when confined to the political sphere seems to be captured in factor 3, which loads strongly on "public opinion polls make the country more democratic," and favorability to the notion of a regular civic poll, along with agreement that "opinion polls do no harm." Factor 4 is a one-item factor, loading .535 on the feeling that "government pays too much attention to opinion polls" and less than .3 on all other items.

The presence of replicate data from 1982 and 1985 makes a stability check on the above interpretations possible. Correlating the factor loadings for 1982 and 1985 (random portion), a method suggested in Harman (1976:343), yields $r = .78$ for factor 1 and $r = .51$ for all four loadings. We gain confidence in the first factor, and become aware that the details of the fourfold structure are vulnerable to instability.

The attitude factors were next entered, along with other variables

Table 3. Response History, Waterloo Survey Respondents

	Yes		No		Total	
Item	*N*	*(%)*	*N*	*(%)*	*N*	*(%)*
To begin with, has an interviewer ever before called at the door to do a survey—except for the census?	62	(36)	110	(64)	172	(100)
IF YES—The last time an interviewer called, did you consent to be interviewed?	53	(85)	9	(15)	62	(100)
Have you ever had a questionnaire arrive in the mail?	125	(73)	47	(27)	172	(100)
IF YES—The last time you received a mail questionnaire, did you fill it out and return it?	80	(64)	45	(36)	125	(100)
Have you ever before been telephoned to be surveyed?	108	(63)	64	(37)	172	(100)
IF YES—The last time you were telephoned for a survey, did you consent to be interviewed?	65	(60)	43	(40)	108	(100)

described below, as predictors of survey response history. The response history differentiated by form of contact, first inquiring, "Has an interviewer ever before called at the door to do a survey—except for the census?" Those reporting a request for a personal interview were then asked, "The last time an interviewer called, did you consent to be interviewed?" The same sequence then ensued for mailed questionnaires and for telephone surveys. A summary of the response history tabulations appears in Table 3. Thirty-six percent of respondents recalled a previous request for an interview at the door, against 73 percent who remembered a mailed questionnaire and 63 percent reporting a telephone contact.[6] The selection bias involved in working with respondents on a survey on surveys prohibits the drawing of any definitive conclusions about the reported response behavior on the different forms of contact. From independent knowledge (Groves and Kahn, 1979:63–76), however, it is known that telephone surveys tend to achieve lower response than personal interviews. That reality is accurately mirrored in the 85 percent responding to the last request for a personal interview, against 60 percent recalling having answered the last telephone survey. The 64 percent response figure for mailed questionnaires is plausible, assuming reasonably vigorous researcher input (Heberlein and Baumgartner, 1978).

Given that refusal probabilities correlate with form of contact, it would be useful to analyze predictors of response for each contact type separately. While we have in unreported analysis examined each component of the history (ensuring, for example, that the sign for the correlation of response history with attitude is consistent across all the tests) it is hard to feel confident about the minutiae of so detailed an investigation of so slim a data base.[7] For personal interview response, after all, the base shrinks to a mere 62 cases. Preferable was the construction of a variable summarizing the response history. Almost half the sample, 77 cases or 45 percent had, judging from their reports, refused *either* the last request for a personal interview, the last mailed questionnaire, or the last telephone contact. This information was entered into a binary variable, contrasting the refusers against the 79 cases who had never refused a survey. For comparability with Table 1, refusers are coded 0, no refusals 1. The number of requests was retained as an independent variable to be entered into analysis later. The 16 cases unable to recall a request for any of the three types of survey described in the response history were deleted.

---

[6] Forty-six percent (random sample) had never been asked for either a personal interview or a telephone survey. That figure compares closely with the 41 percent reported for a spring 1985 survey by Roper, using a different response history question (Roper Organization, 1985:17).

Logistic regression analysis of the response history summary variable appears in Table 4. Since logistic regression results convey limited face value meaning, the emphasis centers upon examining how coefficients for predictors of response alter under different modeling combinations (detected by reading across rows). Starting point, as in Table 1, is the zero-order case where predictors are entered singly. All the attitude variables correlate in the expected direction with response (unfavorable attitudes correspond to high probability of a refusal), with factor 1 exceeding the .05 level of significance, and factor 2 attaining marginal ($p < .1$) significance. Since this result corroborates our earlier judgment that factor 1 captures the most generalized dimension of attitudes toward surveys, this factor advances into the multivariate modeling.

The other variables entered into the zero-order modeling are SES (including the dummy for no information), age, refusal on city directory, sample (1982 vs. 1985), and number of contacts reported on the response history. SES and age merit inclusion since these emerged from Table 1 as the two most serious sociodemographic biases when comparing respondents (given contact) to nonrespondents in the Waterloo surveys. The importance of refusal on city directory, as well as number of contacts on the response history, has been discussed already. Sample is included as a check on how accurately the response histories reproduce across the two samples.

The zero-order coefficients for the nonattitudinal predictors reveal,

---

[7] There is a marginally significant tendency for those interviewed by telephone (i.e., the 1985 Waterloo sample) to hold a higher tendency to have answered the last request for a telephone interview. That relationship is intuitively so sensible it is tempting to believe that it would reproduce in other data sets.

Table 4. Response History Analysis, Waterloo Survey Respondents

	Logistic Regression Coefficients (Dependent Variable = Proportion with No Refusals)				
		Multivariate			
Predictors	*Zero-order*	*1*	*2*	*3*	*4*
Intercept	—[a]	1.773	1.612	1.682	2.430**
Factor 1	.566**		.567**	.494**	.378
Factor 2	.368*				
Factor 3	†				
Factor 4	.236				
SES	−.012	−.012	−.011	−.010	†
No data, SES	†	.522	†	.422	.326
Age	−.019	−.023*	−.022	−.024*	−.023*
City directory refuser	−.916*	−1.484**		−1.316**	−1.205*
Sample	†				
Number of contacts	−.812***				−.714***

[a] See notes, Table 1.

Factor scores: by regression method, from Table 3/Sample: 1982 sample = 1, 1985 sample = 0/Contacts: summed from question sequence described in Table 3, range is 1 to 3/other predictors, see notes to Table 1.

N = 156 (the number reporting at least one prior contact).

first, virtually no ($p > .1$) effect of SES on response history. The effect of age takes the expected (from modeling in Table 1) direction that older people have higher probability of refusal, but fails significance as a zero-order effect. Refusers on the city directory carry high, but only marginally significant, probability of having refused at least one survey request. The next zero-order coefficient confirms that the response histories reproduce accurately across samples, in the sense that no variation in probability of having refused a survey is detectable across the two Waterloo samples. Finally, the zero-order coefficients underline the obvious reality that the greater the number of times a person is asked to participate in a survey the greater the odds that they will have refused at least one request.

The first multivariate model disentangles the effects of city directory refusal and noninformation on occupation. The two correlate, because the city directory by definition can provide no information on a refuser. A suppression is evident, with the marginally significant city directory refuser zero-order coefficient of .916 climbing to a .05 significant level of 1.484. Age now attains marginal (.1) significance, and maintains a marginal probability in most subsequent models.

The second multivariate model tests the extent to which the effect of attitudes toward surveys on response history is independent of sociodemographic factors. The fact that the net effect coefficient for factor 1 is unchanged until the third decimal place (.567 vs. .566) over the corresponding zero-order effect demonstrates that attitude stands virtually orthogonal to the sociodemographic variables.

Next (third multivariate model), city directory refusal reenters the modeling, giving our test of whether attitudes toward surveys are less favorable among those defined from external evidence as survey resistors. This likelihood is confirmed in the test since, sharing variance, the coefficients for both factor 1 and city directory refusal decline (comparing multivariate model 3 with models 1 and 2).

Finally, number of requests for survey response joins the equation, in multivariate model 4. While the zero-order effect that repeated requests enhance the probability of at least one refusal is mathematically self-evident, the multivariate relationship is not. If repeated requests for survey cooperation generated increasingly favorable attitudes to surveying, then controlling for number of requests would enhance the net strength of attitude as a predictor of response history. (Number of requests would suppress, correlating positively with probability of refusal, and negatively with *un*favorable attitude, while unfavorable attitude correlates positively with probability of refusal). But as it turns out, the greater the number of requests for an interview, the greater the unfavorability of attitude (as a Pearsonian correlation, $r = -.22$, $p = .003$), and for this reason controlling for contacts erodes the importance of attitudes, in multivariate model 4. Factor 1 falls below .1 significance, although its positive sign remains reproducible across each sample. The meaning of this and other results receives attention below.

CONCLUSION

The article began from the fact that while the volume and quality of survey on surveys data are increasing, so are reservations about the validity of such information. The central doubts have been that (1) since some appreciable incidence of nonresponse seems unavoidable on any noncompulsory modern survey, respondents on a survey on surveys may be self-selected from that segment of the public most favorably disposed toward the survey method; (2) there may simply be no attitude toward surveying, in the enduring, psychologically deep-seated, reproducible sense.

Our design entailed working at the municipal level, since here external information on sample listings is readily accessible. An analysis of nonresponse bias on surveys on surveys collected in Waterloo, Ontario in 1982 and 1985 revealed a socioeconomic profile not unlike those uncovered in most other surveys (e.g., Hawkins, 1975; Fitzgerald and Fuller, 1982)—an overselection of upper socioeconomic status, hints of low response among such social marginals as the aged and ethnoreligious minorities, and evidence that any correlate of accessibility will prove important, unless response is redefined to comprise no more than the result of confirmed contact.

On the more pointed question, "Do surveys on surveys self-select from believers?" evidence is also at hand: the variable identifying refusal on the city directory—our external measure of orientation toward surveying—performs rather more strongly as a correlate of reported refusal on *other* surveys than in predicting refusal on the survey on surveys. A survey organized around the topic of "What do people think of surveys?" seems to appeal to those opposed to surveying, as well as to the "professional respondents" who thrive on being interviewed on any topic. Thus, it is the elasticity of survey response according to topic, a theme so often stressed in the survey methodology literature (e.g., Heberlein and Baumgartner, 1978:449), that saves surveys on surveys from terminal selection bias.

The analysis of attitudes toward surveying began with a factor analysis of 12 items (some in turn summarizing subitems) from the Waterloo surveys. Four factors were identified by the factor analysis program, with the first conveying the most general meaning when factor loadings were interpreted at face value. The replicative design of the two Waterloo surveys permitted a test of factor stability. Since the surveys were collected three years apart, passing events such as an inaccurate election polling result can be dismissed as common determinants of the factor structure for the two samples. The first factor in the replication correlates highly across samples, enhancing confidence in the reality of the general attitude dimension detected by the factoring. Reproducibility for the secondary factors was not so strong.

The attitude dimensions were then assessed in a construct validity sense, from the argument that if attitudes toward surveying are important, they should correlate with response behavior. The attitude factor variables passed this test, in the believable pattern that the more general the factor the stronger the association with response history. It was discovered in addition that attitudes are related to exposure to surveying, in that the more times a person reports requests for survey cooperation, the more unfavorable is his or her attitude toward the method. It will be useful to see if other researchers can reproduce this potentially important pattern. If the relationship does generalize beyond two small samples collected in a mid-sized Canadian town, surveyors will wish to exercise renewed vigilance against oversurveying and against bogus surveys.[8]

The response history admittedly amounts to the least secure component of the design, since the history itself requires a validity check. Judging from the evidence presented above (note 6), response histories are easily reproducible across alternate question designs, but that does not eliminate the chance of recall errors, conscious and unconscious, in any such question.

---

[8] The depressant effect of contacts on response is not solely due to bogus surveying. The Waterloo respondents were asked: "Has a salesman ever posed as a surveyor in order to try to sell you something?" Including that variable in the computation by no means erases the relationship described in Table 4, model 4.

The findings (Table 4) of a noncorrelation between SES and prior refusal history illustrates the dilemma. Presumably the expected finding fails because the SES effect has already been exhausted by SES-biased selection into the sample of those consenting to participate in the Waterloo surveys. The more worrying possibility, however, is that the response history is so error-prone that a correlation between SES and response history is obscured. Given the low correlation between attitudes toward surveys and SES, the former interpretation seems the more plausible, but it would be satisfying to achieve external confirmation of the accuracy of the response history reports.

Taken as a whole, we believe the evidence presented here belies deeply pessimistic views about surveys on surveys. Indeed, evidence apart, the one part of the pessimistic account negates the other If only surface opinions about surveying exist in the public mind one can hardly speak of professional respondents, and vice versa. Read with due caution and with maximum recourse to external evidence to surmount the epistemological contradiction, surveys on surveys offer potentialities for the market researcher interested in industry image as well as to the academic interested in surveys as social phenomena.

References

Allen, Irving, and J. David Colfax
1968 "Respondents' attitudes toward legitimate surveys in four cities." Journal of Marketing Research 5:431–33.
Arnold, Rome G.
1964 "The interview in jeopardy: a problem in public relations." Public Opinion Quarterly 28:119–23.
Baxter, Richard
1964 "An inquiry into the misuse of the survey technique by sales solicitors." Public Opinion Quarterly 28:124–34.
Biel, Alexander L.
1967 "Abuses of survey research techniques: the phony interview." Public Opinion Quarterly 31:298.
Blishen, B. R., and H. A. McRoberts
1975 "A revised socioeconomic index for occupations in Canada." Canadian Review of Sociology and Anthropology 12:71–79.
Day, George S.
1975 "The threats to marketing research." Journal of Marketing Research 12:462–67.
DeMaio, T. J.
1980 "Refusals: who, where, and why?" Public Opinion Quarterly 44:223–33.
Fitzgerald, Robert, and Linda Fuller
1982 "I hear you knocking but you can't come in: the effects of reluctant respondents and refusers on sample survey estimates." Sociological Methods and Research 11:3–32.

Goudy, Willis J.
1977 Nonresponse Effects: Studies of the Failure of Potential Respondents to Reply to Survey Instruments. Monticello: Council of Planning Librarians.
Goyder, John
1985a "Surveys on surveys: limitations and potentialities." Paper presented at the 40th Annual AAPOR Conference, McAfee, New Jersey.
1985b "Polls on polls show mixed views on polls." Kitchener-Waterloo Record, May 17:A7.
Groves, R. M.
1979 "Actors and questions in telephone and personal interview surveys." Public Opinion Quarterly 43:190–205.
Groves, R. M., and R. L. Kahn
1979 Surveys by Telephone: A National Comparison with Personal Interviews. New York: Academic Press.
Hawkins, Darnell F.
1975 "Estimation of nonresponse bias." Sociological Methods and Research 3:461–88.
Hanushek, E., and J. Jackson
1977 Statistical Methods for the Social Sciences. New York: Academic Press.
Harman, Harry H.
1976 Modern Factor Analysis, 3rd ed. Chicago: University of Chicago Press.
Heberlein, Thomas A., and Robert Baumgartner
1978 "Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature." American Sociological Review 43:447–62.
Kohut, Andrew
1986 "Rating the polls: the views of media elites and the general public." Public Opinion Quarterly 50:1–9.
National Research Council
1979 Privacy and Confidentiality as Factors in Survey Response. Washington, D.C.: National Academy of Science.
New York Times
1985 "Pollsters cite surveys indicating confidence in their work." May 20:B7.
Rogers, T. F.
1976 "Interviews by telephone and in person: quality of responses and field performance." Public Opinion Quarterly 40:51–65.
Rokeach, Milton
1968 "Attitudes." Pp. 449–57 in David L. Sills (ed.), International Encyclopedia of the Social Sciences. Volume I. New York: Macmillan and Free Press.
Roper, Burns W.
1986 "Evaluating polls with poll data." Public Opinion Quarterly 50:10–16.
Roper Organization
1985 Roper Reports, March/April.
Rugg, W. Donald
1972 "Interviewer opinion on the 'salesman as interviewer' problem." Public Opinion Quarterly 35:625–26.
Schleifer, Stephen
1985 "Trends in attitudes toward and participation in survey research." Public Opinion Quarterly 50:17–26.
Schwartz, Alvin
1964 "Interviewing and the public." Public Opinion Quarterly 28:135–42.
Sharp, Laure M., and Joanne Frankel
1983 "Respondent burden: a test of some common assumptions." Public Opinion Quarterly 47:36–53.
Singer, Eleanor
1984 "Public reactions to some ethical issues of social research: attitudes and behavior." Journal of Consumer Research 11:501–09.
Sjoberg, G.
1955 "A questionnaire on questionnaires." Public Opinion Quarterly 18:423–27.

Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s

ROBERT Y. SHAPIRO AND HARPREET MAHAJAN

THE "GENDER GAP" in the public's support for the first Reagan administration made political analysts more attentive than in the past to sex differences in political attitudes and behavior and to the long-term impact of the women's movement. Since the early days of survey research such differences were (or were thought to be) small and inconsequential. Men and women tended to exhibit similar aggregate patterns and trends in voting and political partisanship, although women tended to be less politically active and interested in politics.

Abstract Using 267 repeated policy questions (962 time points), we examine gender differences in policy choices and how they have changed from the 1960s to the 1980s. The average gender difference in preferences toward policies involving the use of force have consistently been moderately large. Sex differences in opinion toward other policies—regulation and public protection, "compassion" issues, traditional values—have been approximately half as large but they also warrant more attention than in the past. Our analysis suggests that the salience of issues has increased greatly for women, and as a result differences in preferences have increased in ways consistent with the interests of women and the intentions of the women's movement.

Robert Y. Shapiro is Assistant Professor of Political Science at Columbia University and a Research Associate at NORC, University of Chicago. Harpreet Mahajan is a Ph.D. candidate in Political Science at Columbia University. The authors owe special thanks to Kurt Veith for his collaboration in some of the research reported here. They are also grateful to Ethel Klein, Eleanor Singer, Ester Fuchs, Eric Smith, Tom Smith, Chaim Eyal, Benjamin Page, Garth Taylor, and the anonymous reviewers for their helpful comments and suggestions. The survey data and tabulations were obtained from the files of the Policy Change and Public Opinion Change project at NORC, which was originally supported by National Science Foundation Grant No. SES-7912969. Support for the research reported here was provided by the Columbia Council for Research in the Social Sciences; the responsibility for analysis and interpretation is the authors'. An earlier version of this article was presented at the 1984 Annual Meeting of the Midwest Association for Public Opinion Research, Chicago, Illinois.

Public Opinion Quarterly Vol. 50:42–61 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-42/$2.50

Significant changes, however, occurred. Not only were women less supportive than men of President Reagan but they had become less likely to consider themselves Republicans (e.g., Clymer, 1983; Schneider, 1984; Frankovic, 1982), and they had become more active politically than in the past and voted in greater numbers (see Wolfinger and Rosenstone, 1980; Verba and Nie, 1972; Welch, 1977; Klein, 1984; Andersen, 1975; Rapoport, 1981). Despite premature reports in November 1984 of a decrease in the gender gap since 1980, the evidence indicates that the sex difference in President Reagan's support was about the same in both elections and that the gender gap may have affected the outcomes of races for other offices (see Opinion Roundup, 1985:28–29; Klein, 1985; Keeter, 1985).

Less attention has been paid to gender differences in *policy preferences* than to voting. Public opinion surveys usually revealed no appreciable differences in choices for government policies, and these seemingly inconsequential findings were not extensively reported and summarized. They were mentioned, however, in voting and public opinion studies (Erikson et al., 1980:186; Monroe, 1975:96; Pomper, 1975; Campbell et al., 1960; Berelson et al., 1954), and breakdowns of responses by sex were often (and are still) routinely reported with survey results and in special compilations, such as those prepared for *Public Opinion Quarterly* (e.g., "The Polls" and the earlier "Quarterly Polls"; see also Cantril with Strunk, 1951).

Not surprisingly, there has been increasing interest in these differences in policy preferences because of their potential electoral importance.[1] Moreover, although some policy issues may be electorally more important than others, any changes that occur in the aggregate distributions of public preferences may ultimately influence the political agenda and the direction of public policy (see Monroe, 1979; Page and Shapiro, 1983). These opinions, to be sure, are not autonomous but may be shaped by events and objective conditions, and by processes of education, leadership, or manipulation (cf. Page and Shapiro, 1982, 1984). Gender and other differences may be important because some groups may be more politically powerful than others or their preferences may be the driving force behind changes in *mass* opinion. Gender differences may take on additional significance because of women's rising levels of political participation (and women simply constituting a large pool of eligible voters) and because of the greater attention they may pay to particular issues and government policies. As a result, even small sex differences in policy choices may no longer be inconsequential, especially if they occur across a wide array of issues around which coalitions can form.

---

[1] For example, Frankovic (1982) showed that women's lower level of support for Reagan could be explained by their greater opposition than men to belligerent positions on war and peace issues, arms control, and the uses of force in foreign policy. This assumes, of course, that these issues and any others were equally important influences on voting decisions for men and women. But this need not have been the case. Klein (1984) showed that some policy preferences for which there were no significant sex differences had different effects on candidate choices. Most notably, men and women had similar opinions toward the ERA and other policies pertaining to women's rights, but over time women appeared to be more likely to take these issues into account in voting. Thus the political importance of policy choices may go beyond the differences and similarities in the positions held by different people. On the other hand, such differences are most likely to occur on salient and important issues, and we would expect larger differences to produce larger effects. Moreover, it is also much more difficult for candidates and policy makers to determine the relative importance of different issues for different groups, *everything else being equal* (i.e., positions on other issues, partisanship; see Klein, 1984; Shabad and Andersen, 1979), than it is to judge simply the extent of differences in preferences.

What do we know about these gender differences? What differences have there been and what changes, if any, have occurred? We will briefly summarize the findings of past studies and then build upon them with an analysis of a larger set of additional data.

Previous Studies

Until recently there were no large collections and analyses of gender differences in various types of political attitudes and opinions. Previous studies examined such differences using a small number of surveys and opinion items, but their results did not lead to much other research or the need to synthesize a broad range of findings. The most thorough analyses were done by Smith (1984), the editors of *Public Opinion* magazine (Opinion Roundup, 1982), Schneider (1984), and Erskine (1971).

Smith (1984) reported that there have been rather striking and persistent gender differences in perceptions and opinions on matters concerned with force and violence, both on the domestic front and in our country's dealings with the rest of the world. Although he did not exclusively examine policy preferences, the differences that he found were compelling: men were more likely than women to support violent options in 87 percent of his readings, and the average gender difference was more than 9 percentage points, with individual differences ranging up to 30 percent.

The *Public Opinion* editors help fill out the picture about other issues by summarizing a sizable but undisclosed number of trends and patterns of differences since 1948. Their findings on violence are consistent with Smith's, except that the gender differences had increased since the mid 1970s. Moreover, they observed that differences had emerged or increased on "compassion" issues concerned with policies which aid the poor, the unemployed, the sick, and others in need, or which protect the well-being of people in general. Women were more supportive than men of such policies (cf. Cook, 1979).

Various authors have explained these emergent differences both by the way women have been socialized as caretakers and nurturers (e.g., see Gilligan, 1982; Stoper and Johnson, 1977; Chodorow, 1974) and by the rise of a women's political consciousness. Women may have extended the scope of their more traditional roles to encompass the problems of society as a whole (see Smeal, 1984; Klein, 1984; Friedan and Decter, 1982). The issues of concern to women thus might include policies to help those in need and also a somewhat distinct set of regulatory policies (e.g., environmental regulation, consumer protection, other safety issues) which are intended to protect the public. The increase in support for these policies would be more likely to occur among newer cohorts of women who, in contrast to earlier generations, came of political age after the emergence of the women's movement. Of course, some women may support the same policies not because of some new political consciousness but because of psychological and developmental differences between the sexes, which lead them to respond in certain ways to pressing new issues. That is, the new issues themselves may be the driving force, but these motivations are difficult to untangle.

Some of these same explanations can be applied to gender differences on force and violence issues and to opinions about policies that preserve and protect traditional values, the home, the family, children, religion, the neighborhood or immediate community, and the order and stability of everyday life. Deviant behavior and personal excesses were to be discouraged and punished. This support for traditional values, broadly defined, is a form of social conservatism which can be compared and contrasted with women's economic liberalism on compassion issues.[2]

---

[2] A single, broadly defined liberal–conservative ideological continuum almost certainly cannot be applied here to women, if it ever could (cf. Nie et al. 1979; Miller and Levitan, 1976; Converse, 1964). It is possible that men and women—or different generations of them—organize their attitudes and opinions in different ways. Compassion, traditional values, regulation and protection, and force and violence may be the major dimensions which help structure women's attitudes. Others may apply to men. The evidence thus far about this type of gender difference (using factor analysis) has been limited and inconclusive; it is uncertain whether men and women systematically organize their political views and policy choices in the same way, if at all (see Goertzel, 1983). What we think is likely is that the dimensions described above do not all apply to all women, but that different ones are important to different women. Thus socialization processes and the consciousness associated with the women's movement can have quite varied influences on political attitudes and opinions.

Some women may perceive women's rights issues themselves, and especially abortion, as a challenge to traditional values. For other women these issues may be distinct, creating only minor tensions with prior values. According to the *Public Opinion* editors, gender differences on traditional values issues appeared to be diminishing, and on women's rights issues there were generally no differences and no changes over time.[3]

Erskine (1971) examined sex differences in attitudes toward women's roles in politics and society (all not policy issues, strictly speaking) for the period from the mid 1930s to the 1970s. She found that prior to the 1960s women were much more supportive than men of expanding women's participation in economic and political life (e.g., employment opportunities, holding public office), although by the 1960s such differences had diminished or even reversed (Erskine, 1971:275).

Since so few studies have been done and since only Smith and Erskine provided statistical estimates of gender differences, many questions remain. We will examine the magnitudes and directions of gender differences for the types of policy issues which we have described, and will appraise the changes that occurred from the 1960s to the early 1980s. By this period women had shifted away from the Republican Party and remained proportionately less Republican than men (Schneider, 1984:1242). These were also the years in which the women's movement first had an opportunity to influence political attitudes and behavior.

Data and Methods

To examine gender differences in preferences we will compare responses to policy questions which have been asked in national surveys. The opinion items that we report here are from surveys conducted by NORC (NORC General Social Surveys (GSS), 1972 through 1983; Davis, 1982 and 1983); Gallup (the 1964–1983 volumes of the *Gallup Opinion Index/Report*); Louis Harris and Associates (from press releases and the *Harris Survey Yearbooks* for 1970, 1971, 1972, and 1973); the Opinion Research Corporation, surveys done for the Advisory Commission on Intergovernmental Relations from 1973 to 1982 (1975–1982), and the Survey Research Center/Center for Political Studies, including the 1952–1978 American National Election Studies (Miller et al., 1980).

---

[3] Cf. Schneider (1984:1243), who states that women were less supportive than men of women's rights 10 years ago, but that the difference has disappeared.

The tabulations from these surveys are to our knowledge the largest set of separate opinion trends for men and women yet assembled. The data consist of 267 different repeated policy questions, and they cover a broad range of foreign and domestic policies. Approximately half contain three or more time points, and they total 962 time points in all for examining gender differences. The data are not a sample, strictly speaking, but they deal with a wide variety of laws, different types of government spending, regulations, acts of government officials, and outcomes, although there are disproportionately more domestic (210 items, 780 time points) than foreign affairs questions (57 items, 182 time points).[4]

One substantive and methodological issue warranted very special attention: how to treat "don't know," "no opinion," and "not sure" responses? We did our analysis both including and excluding them, and we found that the different treatments of don't knows did not affect our results. Past research and some compelling theoretical considerations, however, suggested that the don't knows were important in their own right. Women have always had higher proportions of these responses than men, but there is evidence that this is much less so for new generations of women (see Smith, 1984; Rapoport, 1982). In addition, Shapiro (1982) reported that the proportion of don't knows in general have decreased over time. This conceivably can be attributed in part to the drop in survey response rates, perhaps of apathetic and unattentive members of the public, or to the public getting used to surveys and feeling obligated to provide answers. But the don't knows are also a crude measure of the salience and importance of political issues (cf. Page and Shapiro, 1983; Smith, 1981). The decline in these responses may indicate a rise in the visibility and significance (both personal and public) of policy issues, especially new ones such as the social issues which emerged in the 1960s, and this may have been facilitated by rising levels of education and increases in the amount of information that was easily accessible to the public through television and other media (see Nie et al., 1979). Changes in policy preferences among women may be related to the number of don't know responses, which are affected by attentiveness to politics and interest in issues. Consequently, we estimated gender differences in don't knows and how they changed since the 1960s.

---

[4] The data provide a way of examining the reliability and robustness of the other authors' findings. Some of the items which they used are contained in our data set (e.g., perhaps as many as half of our 164 cases dealing with force and violence were analyzed by Smith (1984)), but there are probably more than 700 additional time points as well.

Analysis and Findings

FORCE AND VIOLENCE

The first question we examined was how different were issues involving the use of force or violence from all other issues. We classified each survey question according to whether it (1) clearly dealt with force or violence, (2) clearly did not, or (3) was ambiguous. For example, issues such as defense, troop levels abroad, capital punishment, and gun control fell into the first category, whereas the space program, the country's activeness in international affairs, the courts' treatment of criminals, and abortion were in the third. As would be expected, the largest number of survey items were in the second category.[5]

Within each category, we calculated the average gender difference (percentage point difference), and we also determined the proportion of time points for which individual gender differences (each time point for a time series) were a substantively and statistically significant (at better than the .05 level) 6 percentage points or more.[6] In one opinion series on abortion, for example, we found 10 successive readings during the period from 1972 to 1983 in the NORC GSS (there were no 1979

---

[5] We were aware that here and later in our analysis there would be some temptation to reclassify issues based on afterthoughts and patterns of findings, especially for policies which tapped several important dimensions at the same time. Doing this leads to larger gender differences for certain categories of issues and strengthens our findings and conclusions. We were uncomfortable, however, with altering our original groupings of issues after-the-fact, since our initial judgments had at least equal substantive merit. Moreover, if the data supported certain theoretical arguments, then the findings should be robust enough to weather the difficulties encountered in categorizing a small number of issues. We leave any reinterpretations of our data to the reader.

[6] To determine the size of each gender difference, we dichotomized the response categories to find the largest difference, making sure, however, in questions with three or more ordinal categories that the *balance* of opinion was different (i.e., that opinion was not merely more polarized for women than men, or the reverse). For the SRC/CPS items which we took from Miller et al. (1980), we divided the reported Percentage Differences Indexes (PDI's) by 2 to make the differences in PDI's roughly comparable to differences in relative frequencies. Don't Know, No Opinion, and Not Sure responses were excluded in the data reported here, except for the SRC/CPS questions.

We initially thought that summing the differences of the aggregate responses (allowing positive and negative ones within a time series to offset each other), taking the absolute values of these sums, and then averaging these absolute values (dividing by the total number of time points involved), would produce small numbers that would be no more than sampling errors for cases in which there were no differences. That is, we expected minuscule average differences because we thought there would be a lot of small and randomly fluctuating positive and negative differences which would simply cancel each other out and produce smaller and essentially meaningless averages. Although the final sums and averages were relatively small, the individual gender differences within each opinion series did not vary much over time and tended *not* to switch signs. Since these differences persisted, it was clear that they did not merely result from sampling error. Thus, allowing for positive and negative gender differences within a given question produced averages that were meaningful and revealed real gender differences.

Table 1. Gender Differences on Policy Issues

	1A. Average Percentage Point Differences				*1B. Percentage of Differences of at Least 6 Percentage Points*			
	Survey Organization				*Survey Organization*			
Issues	*NORC-GSS*	*Gallup*	*Other*[a]	*Total*	*NORC-GSS*	*Gallup*	*Other*	*Total*
Force/ violence	7.6 (52)	8.9 (70)	7.3 (42)	8.1 (164)	71% (52)	66% (70)	67% (42)	68% (164)
Ambiguous	4.4 (112)	5.1 (93)	3.7 (39)	4.6 (244)	27 (112)	40 (93)	33 (39)	33 (244)
Nonforce/ violence	3.8 (215)	4.5 (166)	3.6 (173)	4.0 (554)	33 (215)	38 (166)	24 (173)	31 (554)
Total	4.5 (379)	5.6 (329)	4.3 (254)	4.8 (962)	36 (379)	44 (329)	32 (254)	38 (962)

[a] Includes HARRIS, SRC/CPS, and ACIR/ORC Surveys (80, 123, & 51 time points, respectively). Don't Know, No Opinion, and Not Sure responses have been excluded here and in all other tables for all surveys, except the SRC/CPS surveys.

and 1981 surveys) which showed that men were more supportive than women (by 5, 10, 2, 1, 6, 4, 3, 2, 4, and 7 percentage points) of abortions for women who did not want to have more children. The 4 percent average difference for these represents a true and persistent difference in a certain direction that may have been as great as 10 percent. This type of data analysis admittedly only provides rough estimates, particularly when we take averages over different policies and survey questions. But, as we will see, it can be quite revealing.

Our first results are shown in Table 1, which reports the findings by survey organization. Table 1A presents the average differences for those issues which involve force, those that do not, and those which are ambiguous. Table 1B lists the percentage of time points in which there were gender differences of at least 6 percentage points. The findings are similar for the different survey "houses," which follow different procedures and ask different questions (see Smith, 1978). The gender differences for force issues averaged 8 percentage points, with two-thirds of them over 6 percent, and as we shall see, this almost completely reflects men selecting more violent options.[7] These differences are twice as large, and the 6 percent changes twice as frequent, as those for nonforce issues. The ambiguous policies fall in between, but they are clearly much more like the nonforce issues, suggesting that rather sharp distinctions on the force–violence dimension may be necessary to produce gender differences.

Given the common wisdom about there having been essentially no sex differences in attitudes, the differences that we found for nonforce issues may be somewhat surprising. Six percent differences occurred in one of every three cases, and the average difference of more than 4

---

[7] Some of the exceptions, which may have reflected genuine sampling errors, were small differences on military matters and on permitting a person to hit another person under certain conditions.

Table 2. Average Percentage Point Differences in Specific Policy Directions (% Women − % Men)

	Survey Organization							
Issues	*NORC-GSS*		*Gallup*		*Other*		*Total*	
Opposition to force-domestic	7.6	(39)	12.3	(26)	7.2	(15)	9.0	(80)
Opposition to force-foreign	6.1	(13)	5.6	(44)	7.1	(27)	6.2	(84)
Support for compassion	3.6	(65)	2.0	(40)	3.8	(68)	3.3	(173)
Support for regulation, protection	1.0	(9)	10.0	(21)	2.8	(12)	5.8	(42)
Support for traditional values	3.0	(133)	1.8	(27)	0.0	(39)	2.3	(199)
Support for women's equality, abortion	−1.6	(76)	−3.3	(55)	−1.9	(24)	−2.2	(155)
Other issues (absolute magnitude)	7.9	(44)	3.9	(116)	3.7	(69)	4.6	(229)
Total (absolute magnitude)	3.9	(379)	4.7	(329)	3.5	(254)		

percentage points may not be trivial, if it reflects consistent directions of gender differences and sizable differences for some issues.

The directions of these differences are captured in Table 2. The larger differences for the Gallup domestic-force issues can be attributed to gun control and capital punishment trends for which there are substantial differences.[8] The differences are smaller on foreign and defense matters, although these include few questions on the "gender gap" issues of missile systems and relations with the Soviets. In addition, questions have been included on expanding civil defense (e.g., bomb shelters), which women support more than men by 10 percent or more, despite the possible argument that such defenses might encourage the use of force. Our best judgment is that the gender differences are probably about the same for foreign versus domestic violence, with differences for issues such as gun control, capital punishment, defense spending, the withdrawal of troops from Vietnam, mining the harbors of North Vietnam, and providing arms to the Israelis and the Arabs as high as 15 to 20 percent.[9]

---

[8] E.g., in response to a 1981 Gallup question ("In general do you feel that the laws covering the sale of handguns should be made more strict, less strict, or kept as they are now?") 75 percent of the women respondents versus 56 percent of the men favored stricter laws. Also in 1981, 51 percent of the men versus 38 percent of the women favored capital punishment for treason ("Are you in favor of the death penalty for persons convicted of treason?"). See Smith (1984).

COMPASSION ISSUES

The compassion issues are largely about jobs, income redistribution, and other economic policies to help the poor and targeted group. These include spending on social welfare, education, health, and programs to assist blacks and poor states and central cities, and inflation and unemployment policies.[10] The average gender difference for these policies is just over 3 percent, less than half that of the force issues. The difference is less for the Gallup cases (lowering the overall average), and the issues responsible for this deal with balancing the federal budget: women tended to support a balanced budget by as much as 8 to 10 percent more than do men, despite the possible implications of this for programs that benefit the poor, the elderly, and the like.[11] But women were more supportive of a guaranteed annual income, wage-price controls, equalizing wealth, guaranteeing jobs, government-provided health care, student loans, and rationing to deal with scarce goods.[12]

REGULATION AND PROTECTION

There are larger differences for policies which regulate and protect consumers, citizens, and the environment. The NORC-GSS revealed no difference for spending on the environment, although here the overall average may conceal a change which occurred: by the 1980s perhaps 5 percent more women than men supported spending.[13] In addition, the SRC/CPS and ACIR/ORC surveys include items on the government's general power to protect citizens, perhaps bringing to mind images of "big government" which should be opposed while individual safety regulations are supported. For example, there is overall no appreciable difference in responses to a question about whether the government has too much power,[14] but there are striking differences for particular types of regulations and safeguards. Twenty percent more women than men have expressed opposition to nuclear power plants, and women have also shown more support for the 55 mile per hour speed limit, fines for people who do not wear seat belts, jail terms for drunk drivers, and banning cigarette advertising and sales.[15]

---

[9] E.g., September 1973 Gallup: "There is much discussion as to the amount of money the government in Washington should spend for national defense and military purposes. How do you feel about this: Do you think we are spending too little, too much, or about the right amount?" Percent "too much": 58% women, 45% men. June 1969 Gallup: "Some people think that the time has come to begin to reduce month by month the number of U.S. soldiers in Vietnam. How do you feel—do you think the time has come to do this, or not? Percent "no": 38% men, 21% women. See Smith (1984).

[10] E.g., 1976 NORC-GSS: "We are faced with many problems . . . are we spending too much, too little, or about the right amount on . . .?" Solving the problems of the big cities—percent "too little": 45% men, 51% women. Improving the nation's education system—percent "too little": 49% men, 54% women. Improving the conditions of blacks—percent "too much": 32% men, 23% women. Welfare—percent "too much": 66% men, 60% women.

Taxes were not included in this category of issues; women consistently showed more opposition than men to federal income taxes.

[11] Of course, one may speculate that some women associate this issue with good management or care for the family or household budget.

[12] E.g., January 1969 Gallup: "As you may know, there is talk about giving every family an income of at least $3200 a year, which would be the amount for a family of four. If the family earns less than this, the government would make up the difference. Would you favor or oppose such a plan?" Percent "favor": 31% men, 38% women. September 1980 Gallup: "Would you favor or oppose having the government bring back wage and price controls?" Percent "favor": 52% men, 61% women.

The traditional values issues include pornography, the use of drugs such as marijuana, prohibition, prayer in schools, sex education, providing birth control for teens, and homosexuality and other alleged forms of deviant behavior. Women have consistently tended to uphold conservative values but the overall average difference is small.[16] The average for the NORC-GSS is highest because of the inclusion of several of Stouffer's tolerance/civil liberties items, in response to which women still express more support than men for limiting freedom of speech for would-be threats to society such as Communists, socialists, or atheists (see Stouffer, 1967; Nunn et al., 1978; McClosky and Brill, 1983).[17] We also included civil rights issues in this category, because these questions usually dealt with intermarriage, housing, and schools.

---

[13] See footnote 10 for question wording: " . . Improving and protecting the environment." In 1974, 5 percent more men than women responded (spending) "too little," whereas in 1978, 6 percent more women than men said "too little," 5 percent in 1980, and 2 percent in 1982.

[14] E.g., May 1978 ACIR/ORC: "Which of these statements comes closest to your view about government power today?" Percent "Federal government has too much power": 40% men, 41% women.

[15] E.g., February 1981 Gallup: "Do you favor or oppose keeping the present 55-mile-per-hour speed limit on the highways of the nation." Percent "favor": 72% men, 81% women. June 1981 Gallup: "Do you think there should or should not be a complete ban on cigarette advertising?" Percent "should be": 39% men, 53% women.

[16] E.g., March 1970 Gallup: "Do you think they (birth control pills) should be made available to teenage girls?" Percent "no": 72% men, 87% women. June 1980 Gallup: "Do you think the use of marijuana should be made legal, or not?" Percent "yes": 33% men, 20% women.

[17] E.g., 1982 NORC-GSS: "Now I would like to ask you some questions about a man who admitted he is a communist. Suppose this admitted communist wanted to make a speech in your community should he be allowed to speak, or not?" Percent "yes, allowed to speak": 62% men, 55% women.

Here differences varied depending on the issue. Overall there is no consistent and clear pattern, although if one takes into account compassion issues and positions concerning the general social and economic welfare of minorities, women perhaps tend to be slightly more liberal than men. When the civil rights issues are discounted among the traditional values policies, the average gender difference roughly doubles in the conservative direction.

These values may clash with abortion and with various facets of women's rights and equality. As we see in Table 2, women are overall more conservative on women's rights issues, which is almost completely due to men being consistently more pro-choice on abortion (based on questions concerning support for abortions under varying circumstances, such as rape, potential birth defects, danger to the woman's health). On other women's issues there are few clear and consistent differences.[18]

For the remaining (unclassified) issues Table 2 reports the average absolute magnitude of the differences. The average is 4.5 percent, and there are some large and noteworthy differences, especially among the NORC-GSS questions. For instance, support for wiretapping was more than 10 percentage points higher for men than women (women perhaps seeing this action as a forceful intrusion upon people's homes and lives), and women were more opposed (by over 15 percent) to spending on the space program (perhaps not so much because of its potential role in national defense as because spending on it used resources that might be used for other purposes). Some issues may tap several policy or ideological dimensions that are important to women, but it may be difficult to determine, a priori, their relative importance.[19]

Trends in Gender Differences

Thus far we have confirmed that there are gender differences on the force dimension and other significant variations as well. Women's concerns are a mixture of economic liberalism, government activism, and traditional social conservatism (with some social liberalism as well).

---

[18] This absence of a gender difference, however, obscures the possibly divergent sources of these preferences. The positions that men take on women's issues may be more likely to involve symbolic politics and liberalism, whereas women's choices may be more closely related to their experiences (Klein, 1984). As suggested earlier, men and women may be structuring their attitudes toward these issues in different ways.

[19] Other issues on which gender differences occur are campaign and election reforms, including the lengths of terms of office (with women more in favor of reforms), and policies toward labor and union activities. These differences cannot as readily be accounted for as the others.

Table 3. Changes in Gender Differences for Different Types of Issues (% Women − % Men)

	Time Period							
Issues	*1964–71*		*1972–76*		*1977–83*		*Total*	
Force-domestic	12.2	(14)	8.3	(29)	8.4	(37)	9.0	(80)
Force-foreign	7.2	(23)	5.7	(27)	5.9	(34)	6.2	(84)
Compassion	3.2	(29)	2.2	(55)	4.1	(89)	3.3	(173)
Regulation, protection	3.0	(5)	4.9	(12)	7.2	(25)	5.8	(42)
Traditional values	−0.9	(38)	2.2	(84)	3.6	(77)	2.3	(199)
Women's issues	−3.6	(8)	−2.1	(51)	−2.3	(53)	−2.2	(155)
Other	5.1	(73)	4.5	(71)	4.8	(85)	4.6	(229)
Total (absolute magnitude)	4.6	(190)	3.6	(329)	4.4	(443)		

But did these differences simply persist from the 1960s to the 1980s? The rough trends are presented in Table 3, in which we compare the average differences for three time periods, 1965–71, 1972–76, and 1977–83. In general the changes are modest.[20] The gender differences for domestic-force issues show an average drop of 4 percent from the 1964–71 period to 1977–83. This occurred through a small narrowing of the differences on capital punishment, gun control, and urban unrest, which still remained moderately large. Women may have been responding to higher crime rates, especially those crimes of which women are more likely to be victims (and perhaps increases in gun ownership among women contributed to their shift in opinion against gun control). Foreign-force showed a slight drop (probably distorted by civil defense policies and the omission of the space program and recent foreign and defense issues). There was no change from 1972–76 to 1977–83, and our best judgment is that there has been no change in gender differences on force issues taken as a group.

The remaining issues, except for those unclassified, show more interesting changes. For the compassion issues and more so for regulation–protection we find increases in gender differences. The differences for the latter are as big as those for all the force issues. One explanation for this is that after 1976, fiscal problems and (threats of severe) budget cuts in social spending and proposals for deregulation, despite visible safety and environmental problems, began to have an impact on public opinion, especially during the Reagan administration; and women reacted to this differently than did men.[21] The results are consistent with arguments about women's societal interests, although the table shows that the trends for compassion and regulation issues are not paralleled by the war and peace and other force issues, as some authors might have expected (cf. Smeal, 1984). Nor do the shifts for the compassion and protection issues fully mirror each other. They moved differently from the 1964–71 to the 1972–76 periods, but then shifted in tandem during the subsequent 1977–83 period, which is the crucial one. What may have happened here and in the cases of other issues as well was that the movements in aggregate opinion for different issues were determined by shifts in opinion by *different* women. Any questions about the characteristics of these women need to be answered elsewhere at the individual level, although we can speculate that at the extremes are women committed to ideas associated with the women's movement versus those who staunchly defend more traditional attitudes.

---

[20] For a single trend, a change in subgroup differences of about 10 percentage points might be considered substantively and statistically significant. Shapiro and Page (1984) found that diverging or converging trends of this sort are infrequent and less likely to occur for sex than for other population characteristics. In Table 3, however, we are comparing trends for a *large number of issues*, and smaller but consistent differences for them can be politically important. Cf. Taylor (1980) concerning pooled trends.

This becomes more apparent when we examine the traditional values issues. In contrast to compassion and regulation, women were more conservative on these matters, and the gap appeared to widen. The increase in these differences, however, is largely attributable to the different mix of issues included in this category, especially in the first and last periods. Several Gallup and Harris civil rights trends are in the 1964–71 period, and the NORC-GSS (Stouffer) tolerance items and questions on drugs, alcohol, and pornography are in the later time periods. In contrast, there may have been a slight drop for women's issues, but women still appear to be more conservative due to differences in opinion toward abortion.

Thus, a large number of women came to be even more liberal than men on compassion and regulation issues, while they or other women moved in a conservative direction or did not change their preferences on certain traditional values or social issues. Table 3 summarizes these dynamics but it offers no evidence about who altered their preferences and how this change came about. Women's greater concern than men for certain needy groups and for the well-being of society may have occurred because of their upbringing and other experiences in life. For some women the policies supported by the women's movement may have contributed directly, or the movement's impact may have occurred indirectly as it made politics and issues more salient and important for women than in the past. This may have occurred without having a uniform effect on women's attitudes. Women who wanted traditional values protected, especially women in older cohorts, may have been led to form and express their preferences. Some may have reacted negatively to liberal positions taken by participants in the women's movement which they perceived as extreme.

---

[21] As the public as a whole became more conservative during part of this period, this shift was greater for men than for women.

We would expect new generations of women to behave in a distinct way, or at least that the sources of their attitudes and behavior would be different from earlier cohorts. They are better educated and are more likely than members of past generations to be in the workforce (and for longer periods) and to experience more economic independence. They have been brought up in a society in which the women's movement has had a chance to exert direct or indirect influence. We would anticipate, then, that politics and issues would have become more important than in the past for women, in comparison to men, because of social and political influences on all women or as new cohorts became a larger part of the population. These competing explanations, of course, are not mutually exclusive; several processes can be at work at the same time.

Decline in Don't Know Responses

Rapoport (1982) examined some evidence concerning this. He observed that new cohorts of women gave fewer don't know responses to questions about politics asked in national surveys, and this very likely reflected the greater visibility of issues and politics for women and an increase in the attention that they paid to them. If the changes that we found can be attributed to the greater salience of these issues among women of all ages, or at least among younger cohorts (which we cannot untangle from our aggregated data), we would expect a decrease in the percentage of don't know responses on those issues for which there were changes in gender differences, that is, for compassion issues, regulation and protection, and traditional values.

We found that don't know responses have declined since the 1960s for both men and women (see Shapiro, 1982:80–81). Looking at the data by survey organization, since proportions of don't knows are highly sensitive to "house effects" (Smith, 1978), we find that the Gallup don't knows for men dropped from 8.1 percent (*n* = 110) to 7.3 percent (*n* = 59) and 7.4 percent (*n* = 160) when we compare the three time periods; for women the figures were 12.9, 11.1, and 9.9 percent. For men in the Harris surveys we found 12.2 percent (*n* = 31) for 1964–71 versus 9.2 percent (*n* = 49) for the later periods, and 18.3 to 12.5 percent for women. The ACIR/ORC survey showed 12.1 percent (*n* =

Table 4. Changes in Gender Differences in "Don't Know" Responses for Different Issues (% Women − % Men)

	Time Period							
Issues	*1964–71*		*1972–76*		*1977–83*		*Total*	
Force-domestic	1.7	(11)	1.9	(23)	2.0	(32)	1.9	(69)
Force foreign	5.2	(18)	5.3	(24)	5.7	(32)	5.4	(74)
Compassion	5.8	(17)	3.9	(47)	3.1	(69)	3.7	(133)
Regulation, protection	4.0	(1)	3.8	(10)	2.7	(23)	3.1	(34)
Traditional values	3.4	(24)	0.9	(73)	0.7	(75)	1.2	(172)
Women's issues	−1.3	(7)	−0.2	(45)	0.7	(92)	0.4	(144)
Other	6.8	(63)	4.5	(67)	3.8	(83)	4.9	(213)
Total	5.1	(141)	2.6	(289)	2.4	(409)	2.9	(839)

17) for men in the 1972–76 period versus 9.6 ($n = 34$) for 1977–83, and 18.4 versus 13.6 percent for women. The NORC-GSS showed smaller changes in don't knows.

Issues became more salient for everyone but especially for women, judging from the larger changes in don't knows. One possible reason for the smaller change among men is that they began with fewer don't knows to begin with and that there may be some kind of floor effect. That is, some of these responses may not occur because of the low salience of issues but rather because of the difficulty that politically attuned respondents may have in making their opinion fit one of the categories offered in closed-ended survey questions. These don't knows do not necessarily decline as issues become more important to people, and the don't knows among the men may have been of this type. Women may have only more recently neared the floor level of "educated" don't know (see Smith, 1981). Still, the importance of the decline in don't knows and the increase in the salience of issues among women is not diminished.

Table 4 summarizes the changes in the gender differences in don't knows. Historically there has been about a 3 percent difference, according to what we know from the survey research literature (Smith, 1984:383). That is approximately what we find (2.9 percent), but from the 1960s to the 1980s the difference has declined from 5 percent to approximately half that. The largest difference remains for foreign policy issues (force and several other issues). Not surprisingly, the smallest difference is for women's issues, men often having had more don't knows.

The overall drop in don't knows indicates that women have been catching up in terms of the attention that they pay to specific issues. It is possible that women were always as attentive as the data show for the most recent period, but that they were reluctant to express their views. They may have deferred to the opinions of men, or perhaps the process by which this occurred involved the hesitance (or intimidation) that people may feel if they think that their opinion is in a minority; "pluralistic ignorance" and the "spiral of silence" might have played a part (cf. Noelle-Neumann, 1974). The women's movement may have helped overcome this by legitimating and encouraging women to express their positions whether they were in the majority or not, and by the same token whether or not they supported the political positions of the movement itself.

Alternatively, the gender differences may have appeared to change because men reached the level at which don't know responses are no longer sensitive to changes in the salience of issues. In the 1960s and 1970s issues played a bigger part in *everyone's* lives. It was a dynamic period in American history. Issues such as civil rights, political protest, the Vietnam war, crime and other matters involving social control excited the nation. The media, especially television news, were there to cover the relevant events and public debate and to present their reports to a citizenry that was better educated than in the past (e.g., see Nie et al., 1979).

The findings in Table 4 and other evidence suggest that something more than the appearance of new and salient issues was affecting women. The don't knows declined for men and women on all issues, but the gender differences were more likely to decrease for issues on which the gender differences in preferences increased. Moreover, these were the issues for which opinions changed in directions that were consistent with the kinds of interests and concerns that women—perhaps very likely not the same women—may have already had or developed. For the compassion, regulation, and traditional values issues, the gender differences in don't knows were associated with *increases* in differences in preferences. In contrast, there was no change in the differences in don't knows for the force issues, while there was a decline or no change in the differences in policy choices.

If the floor effect argument provides a sufficient explanation, then the decrease in gender differences in don't knows should have occurred for all issues. It did not occur for the foreign-force issues, even though the large difference at the start suggests that these issues should have been less sensitive to floor effects. And if no policies became more important to women, the decline in don't knows should not have led to larger sex differences on particular issues. Thus, women's psychological involvement with questions of government policy and women's policy choices themselves may have been influenced distinctively by changes in society and politics.

Conclusion

Our analysis improves our understanding of gender differences in policy preferences and how they have changed from the 1960s to the present. During this period there were persistent differences in preferences. The average gender difference in preferences toward policies dealing with the use of force and violence have consistently been moderately large. Sex differences in opinion toward other policies have been approximately half the size, but they also merit more attention than they have been given in the past. The next largest differences occur for regulation and other protective policies and for compassion issues since the mid 1970s, which may be a consequence of social and political changes. The salience of these issues for women increased more than it did for men, and gender differences in opinions increased in ways suggested by the interests which women have had and consistent with the intentions of the women's movement. Similarly, women as a group also supported the protection of conservative traditional values more than did men. The impact of the women's movement is less clear here. On issues such as abortion and women's rights, women's greater conservatism probably diminished, but there are still small but persistent gender difference in opinions toward abortion.

As we get further into the 1980s and the second term of the Reagan administration, we will be able to find out whether the differences reported here remain or whether the trends continue or reverse. Because of ongoing changes in economic, social, and political life in the United States, such differences in policy preferences are likely to be looked at more closely in the future than they have been in the past. And it will be interesting to see how these differences affect electoral politics and the policy-making process.

References

Advisory Commission on Intergovernmental Relations
Changing Public Attitudes on Governments and Taxes. Washington, D.C.: ACIR. 1975, 1976, 1977, 1980, 1981, 1982 Vols.
Andersen, Kristi
1975 "Working women and political participation, 1952–1972." American Journal of Political Science 19:439–55.
Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee
1954 Voting. Chicago: University of Chicago Press.
Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes
1960 The American Voter. New York: Wiley.
Cantril, Hadley, with Mildred Strunk
1951 Public Opinion 1935–1946. Princeton, NJ: Princeton University Press.
Chodorow, Nancy
1974 "Family structure and family personality." In Michelle Zimbalist Rosaldo and Louis Lamphere (eds.), Women, Culture, and Society. Stanford, CA: Stanford University Press.

Clymer, Adam
1983 "If anything, gender gap is becoming even wider." The New York Times. Sunday, December 11, p.E5.
Converse, Philip E.
1964 "The nature of belief systems in mass publics." Pp. 206–61 in David E. Apter (ed.), Ideology and Discontent. New York: The Free Press.
Cook, Fay L.
1979 Who Should Be Helped? Beverly Hills, CA: Sage Publications.
Davis, James A.
1982 &
1983 General Social Surveys, 1972–82 and 1972–83: Cumulative Codebook. NORC, University of Chicago.
Erikson, Robert S., Norman R. Luttbeg, and Kent L. Tedin
1980 American Public Opinion: Its Origins, Content, and Impact. 2nd ed. New York: Wiley.
Erskine, Hazel
1971 "The polls: women's role." Public Opinion Quarterly 35:275–90.
Frankovic, Kathleen
1982 "Sex and politics—new alignments, old issues." PS 15:439–48.
Friedan, Betty, and Midge Dector
1982 "Are women different today?" Public Opinion April/May:20, 41.
Gilligan, Carol
1982 In a Different Voice. Cambridge, MA: Harvard University Press.
Goertzel, Ted George
1983 "The gender gap: sex, family income and political opinions in the early 1980s." Journal of Political and Military Sociology 11:209–22.
Harris, Louis, and Associates, Inc.
The Harris Survey Yearbook of Public Opinion, 1970–1973. 4 vols. New York: Louis Harris and Associates, Inc., 1971, 1975, 1976.
Keeter, Scott
1985 "Public opinion in 1984." Pp. 91–111 in Gerald Pomper et al., The Election of 1984: Reports and Interpretations. Chatham, NJ: Chatham House.
Klein, Ethel
1984 Gender Politics. Cambridge, MA: Harvard University Press.
1985 "The gender gap: different issues, different answers." The Brookings Review 3:33–37.
McClosky, Herbert, and Alida Brill
1983 Dimensions of Tolerance. New York: Russell Sage Foundation.
Miller, Warner E., and Teresa Levitin
1976 Leadership and Change: Presidential Elections from 1952–1976. Cambridge, MA: Winthrop Publishers.
Miller, Warner E., Arthur H. Miller, and Edward J. Schneider
1980 American National Election Studies Data Sourcebook, 1952–1978. Cambridge, MA: Harvard University Press.
Monroe, Alan D.
1975 Public Opinion in America. New York: Dodd, Mead.
1979 "Consistency between public preferences and national policy decisions." American Politics Quarterly 7:3–19.
Nie, Norman H., Sidney Verba, and John R. Petrocik
1979 The Changing American Voter. Enlarged ed. Cambridge, MA: Harvard University Press.
Noelle-Neuman, Elisabeth
1974 "The spiral of silence: a theory of public opinion." Journal of Communication 24:43–51.
Nunn, Clyde Z., Harry J. Crocket, Jr., and J. Allen Williams, Jr.
1978 Tolerance for Nonconformity. San Francisco: Jossey-Bass.
Page, Benjamin I., and Robert Y. Shapiro

1982 "Changes in Americans' policy preferences, 1935–1979." Public Opinion Quarterly 46:24–42.
1983 "Effects of public opinion on policy." American Political Science Review 77:175–90.
1984 "Presidents as opinion leaders: some new evidence." Policy Studies Journal 12:649–61.

Pomper, Gerald M.
1975 Voters' Choice: Varieties of American Electoral Behavior. New York: Dodd, Mead.

Public Opinion
1982 "Opinion roundup: women and men: is a realignment under way?" April/May, pp. 21, 27–32.
1985 "Opinion roundup: III—the exit poll results." December/January, pp. 28–29.

Rapoport, Ronald B.
1981 "The sex gap in political persuading: where the 'structuring principle' works." American Journal of Political Science 25:32–48.
1982 "Sex differences in attitude expression: a generational explanation." Public Opinion Quarterly 46:86–96.

Schneider, William
1984 "Opinion outlook: the democrats are counting on the gender gap, but it may not be much help." National Journal, June 23, 1984, pp. 1242–43.

Shabad, Goldie, and Kristi Anderson
1979 "Candidate evaluations by men and women." Public Opinion Quarterly 43:18–35.

Shapiro, Robert Y.
1982 "The dynamics of public opinion and public policy." Doctoral dissertation, University of Chicago.

Shapiro, Robert Y., and Benjamin I. Page
1984 "Subgroup trends in policy choices: a preliminary report on some theories and findings." Paper presented at the annual meeting of the Midwest Political Science Association.

Smeal, Eleanor
1984 Why and How Women Will Elect the Next President. New York: Harper and Row.

Smith, Tom W.
1978 "In search of house effects: a comparison of responses to various questions by different survey organizations." Public Opinion Quarterly 42:443–63.
1981 "Educated don't knows: an analysis of the relationship between education and item nonresponse." GSS Technical Report No. 28. NORC, University of Chicago.
1984 "The polls: gender and attitudes toward violence." Public Opinion Quarterly 48:384–96.

Stoper, Emily, and Roberta Ann Johnson
1977 "The weaker sex and the better half: the idea of women's moral superiority in the American feminist movement." Polity 10:192–217.

Stouffer, Samuel A.
1967 Communism, Conformity, and Civil Liberties. New York: Wiley. Original 1955.

Taylor, D. Garth.
1980 "Procedures for evaluating trends in public opinion." Public Opinion Quarterly 44:86–100.

Verba, Sidney, and Norman H. Nie
1972 Participation in America. New York: Harper and Row.

Welch, Susan
1977 "Women as political animals." American Journal of Political Science 21:711–30.

Wolfinger, Raymond E., and Steven J. Rosenstone
1980 Who Votes? New Haven, CT: Yale University Press.

Preelection Polls in Israel: Structural Constraints on Accuracy

JACOB SHAMIR

REPORTING preelection polls in the mass media is a popular and widespread practice in Western democracies. Much has been written on the definition of polls as news (e.g., Crespi, 1980; Tuchman, 1978), styles of reporting (Broh, 1980) and standards for accurate presentation of results (Field, 1971, Meyer, 1973; Noelle-Neuman, 1980; Miller and Hurd, 1982). Much of social science research methodology literature, including many of the articles that have appeared in this journal over the years, have devised techniques to improve the accuracy of such polls.

This article approaches the issue of election polls' accuracy from a different angle. It suggests that accuracy of reports on preelection polls is greatly affected by the environement in which pollsters operate as well as by structural characteristics of the system itself. Hennessy and Hennessy (1961) have already found that partisan-sponsored polls overestimate the actual vote for the candidate of the party for which the polls were done. They attribute this effect to pollsters' wish to please their partisan employer. However, their study suffered from serious sampling as well as other methodological problems (Perry,

---

Abstract This article examines the last two Israeli election campaigns of 1981 and 1984, and suggests that structural variables greatly affect the accuracy of reported preelection polls. Pollsters' quest for more detail and refinement driven by competition and mass media pressures for highly processed predictions is liable to result in less accurate reports. The kind of errors in polls' projections are situation-bound and greatly depend on the nature of the floating vote in the system. Interpollster accuracy, however, is mainly determined by pollster–sponsor dependencies. Pollsters working for political parties projected results more favorable to their sponsors or less favorable to their sponsors' opponents. Moreover, the stronger the pollster–sponsor structural dependencies, the greater the inaccuracy.

Jacob Shamir is Lecturer of Communication at the Journalism Program, Tel-Aviv University, Tel-Aviv, Israel.

Public Opinion Quarterly Vol. 50:62–75 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-62/$2.50

1962). I suggest that biasing pressures on pollsters are structurally inherent in the pollster–sponsor work relationship. The closer these relations get and the more structurally dependent pollsters become, the more favorable their reports are to their sponsors.

Structural characteristics of the system itself further affect accuracy and may facilitate such partisan pressures. Some political systems and situations are less transparent than others to preelection predictions. By transparency I mean the amount of practical and methodological difficulties that inherent systemic characteristics impose on pollsters' prediction efforts. A system's transparency depends on such factors as the number and size of parties in the system, the electoral system, and the public's willingness to volunteer valid voting intentions. The quest for higher precision and detail in a low transparency system—driven by competition and mass media pressures for highly processed predictions—is liable to result in less accurate reports, especially when partisan relationships with sponsors are involved. These propositions are examined in the Israeli system in the last two election campaigns of 1981 and 1984; however, they may be generalized to other polling situations and other political systems.

A System's Transparency to Polling Efforts

Israel's political system since the 1977 election, which marked the end of the dominant party system, poses a challenging environment to pollsters attempting to predict election outcomes with accuracy. Deep ideological cleavages coupled with a proportional representation electoral system resulted in two major political blocks almost equal in size. Each is composed of a major party and a host of satellite parties. A smaller third block consists of religious parties.

The abundance of political parties in the Israeli system, many of which are rather small, poses several methodological problems for pollsters. First, predictions with respect to the smaller parties, in particular those on the verge of passing the one percent barrier clause, are poor and risky because of the pollsters' commonly used sample size of 1200 subjects and its resulting sampling error. Second, it is much harder to classify no answer responses and undecided voters among the various political parties with high confidence. Subjects' stand on debated issues and attitudes toward party leaders may serve as powerful predictors in classifying floating voters among the two major blocks but are less valid in classifications among the various parties within a specific block or among other smaller parties. This finer voting distribution is critical in the Israeli context since it may determine the ability of the major parties to form a coalition government.

The major problem for Israeli pollsters, however, is the sizeable number of undecided voters, refusals to reveal voting intentions, and insincere responses in the system (Diskin, 1981). For example, voters of the first two kinds constituted about 30 percent of the population only one week before the 1981 and 1984 elections.[1] When the fractionalized party system interacts with the large number of ambiguous voting responses, the dimensions of the pollsters' problem increase many fold. A complex system of allocating residual Knesset seats among the various parties after the election poses another problem for pollsters attempting to cast their predictions in terms of number of seats.

These last two problems can be avoided if predictions are given in raw percentages rather than Knesset seats and hard to classify votes are left unclassified. The price, of course, is cruder prediction and less meaningful results for decision makers and mass media audience.

Additional problems are posed by two especially hard to reach populations of the Israeli constituency—soldiers and Arabs. These groups are usually not polled because of sampling, interviewing, and cooperation problems.

One may have expected a system such as this to encourage a conservative approach on the part of pollsters in treating the crude data obtained in the field and in the degree of specificity of their predictions. It seems, however, that with the increase in the public's interest in polls there is rather an increase in pollsters' attempts to provide highly processed data and more specific predictions. The number of pollsters casting their predictions in terms of seats won rather than percentages increased monotonically since 1977.

This trend correlates with changes in the amount and style of polls' coverage in the Israeli press. Weimann (1982), for example, found that polls' media coverage peaked at the 1981 election, reaching a state of "obsession for polls." The 1984 campaign was if anything even "worse" in this respect. Pollsters themselves became public figures heavily interviewed in the mass media and challenging each other's predictions.[2] The style of polls' coverage in the media reached the extreme characteristics of what Cantril (1976) and Broh (1980) called "horserace journalism." Thus, most reports on polls in 1984 focused on the size and dynamics of the gap between the two major parties, on the electoral fate of certain smaller extremist new parties, and on "beauty contest" ratings of the candidates for various government posts.

---

[1] Kessafim No. 54, 6/29/1981; Ma'ariv 7/19/1984.

[2] For example, Ma'ariv 7/13/1984, Hadashot 7/18/1984.

Data Refinement Efforts

Pollsters' main data refinement efforts center on developing valid classification methods to handle the great number of hard to classify responses obtained in the field. These methods, designed to refine predictions, may paradoxically become a major source of prediction error.

Commonly, undecided voters are investigated with respect to their acceptable alternative parties. The known alternatives are then used to classify them as either floating between the two major political blocks or within a major block. Then decision rules for final classifications take into account direct and projective attitudinal measures, subjects' declared vote in the last election, and at least some experience and judgment on the part of the pollster. Subjects who refuse to reveal their vote are handled much the same way except that it is more risky to initially classify them into one of the major blocks.

These procedures constitute a major source for classification errors for several reasons. The theoretical foundations of classification rules are quite weak and the boundaries between objective classification criteria and subjective judgment are not always clear, which opens the way to biasing structural effects. Second, these decision rules may also be very situation-bound, performing well in some respects, in one election, and proving to be wrong in another. Third, some of the objective classification measures may be of low validity. For example, subjects' information concerning their last vote may be false (Diskin and Felsenthal, 1981). Some pollsters handle this by using some aggregate data correction procedures which take into account the gap between subjects' declared voting distribution in the last election and the actual voting results.

Other potential sources of error are the assumptions made by pollsters on the vote distribution in the Arab sector. These assumptions are usually based on outdated last election statistics, subjectively adjusted to fit the present situation. As the Arab sector vote accounts for about 10 out of 120 Knesset seats, it may constitute a considerable source of prediction error. Since these data refinement procedures are based at least partly on subjective judgment and experience, it is easy to see how organizational, social, and economic factors may greatly affect their application.

Pollster–Sponsor Structural Dependencies

The last two elections, which varied considerably with respect to the pollster–sponsor relationships, can provide a good test for our propositions. In the 1981 campaign the Likud contracted one of the two largest commercial polling organizations in the country—Modi'in Ezrahi Research Institute (MERI). The Alignment in turn contracted the other one, which was affiliated with the advertising agency working with them—Dahaf. Both political parties, however, commissioned additional polls from other sources as well. Similarly, both Dahaf and MERI did not work exclusively for these political parties but for other parties as well, and provided regular polling services to several media organizations.

In 1984 the two major parties started by contracting commercial polling organizations which were dismissed later in the process in favor of free-lance pollsters operating ad hoc "in-house" polling facilities.[3]

Clearly, then, the Israeli polling industry witnessed considerable change in pollster–sponsor dependencies from the 1981 to the 1984 elections. First, free lance pollsters working exclusively for a political party have a considerably greater economic dependence than do polling organizations serving additional clients as well.[4] However, more important is the physical and social proximity stemming from the use of party facilities and the involvement of interested party officials in polling activities. Close proximity and ongoing meetings between the pollster and the sponsor's representatives are likely to reinforce informal social ties and the transmission of expectations concerning polls' results. Moreover, such conditions further decrease the pollster's control over the content and style of the poll's results, in a system already suffering from lack of self-regulation and public supervision over pollsters' activities.

The next section shows that systemic characteristics, i.e. the number and kind of the floating votes in the 1981 and 1984 Israeli elections, determined to a large extent the kind of errors in predicting election results. It also shows that pollsters working for political sponsors tend to favor their sponsors in preelection predictions. Moreover, inaccuracies in predictions increase with increases in pollsters' dependence on their political sponsors.

Analysis

The data used in the analysis consist of 31 out of 32 polls[5] carried out by Israel's most prominent pollsters and published in the press during

---

[3] The Alignment pollster established and directed a full-scale in-house polling organization, operating within party facilities. The Likud pollster also used party facilities to direct polling efforts, subcontracting outside research organizations for data collection.

[4] The Alignment pollster in 1984 seems to have volunteered, receiving only expenses from the party.

[5] See note c, Table 3.

Table 1. Styles of Reporting Preelection Projections in 1981 Israeli Election Campaign

	IASR[a] *(1st week of June)*			*SRC*[b] *(June 4–11)*	*Dahaf*[c] *(June 1–6)*	*MERI*[d] *(June 2–7)*
Party	*Type of Vote*	*% Projected*	*Party*	*% Projected*	*Seats Projected*	*Seats Projected*
Likud	definite	19	Likud	38	42	46
Alignment	definite	24	Alignment	33	43	40
Alignment or CRM/Shinuy/Shelli/Tami	floating	8[e]	NRP	6	6	8
Alignment or Telem/Tchiya/NRP	floating	3	Aguda		5	5
Alignment or Likud	floating	3	Telem		4	3
Likud or Telem/Tchiya/NRP/Tami	floating	10	Tami		4	3
Other parties	definite/floating	18	Tchiya		3	2
Unclassified	unknown	15[f]	CRM			2
			Other parties	9	13	11
			Undecided	14	15.8	14.5
			Not sure		19.2	

[a] Israel Institute for Applied Social Research.
[b] Smith Research Center.
[c] Dahaf Research Institute.
[d] Modi'in Ezrahi Research Institute.
[e] Indicated vote preference but not sure in their decision.
[f] Did not indicate vote preference (don't know or no answer).

the last two months of the 1981 and 1984 election campaigns. These are the hottest months in terms of political parties' struggle for voters' support and the most crowded with preelection predictions appearing in the mass media.

Most polls were published on a regular basis under various contractual arrangements between the Israeli major newspapers and their pollsters. Some polls, however, in particular the Likud and Alignment polls in 1984, were made public either in press conferences or less formally through political correspondents covering the election campaign.

Let me first illustrate the various pollsters' approaches in reporting their results. Table 1 presents results of four polls carried out by four prominent Israeli polling organizations less than a month before 1981 election day. The results are presented in the same format published in the press.

The first major point evident from the table is the sizeable percentage of respondents floating between at least two alternative parties. IASR, which is a nonprofit research organization affiliated with the Hebrew University in Jerusalem, reports floaters' data in great detail. Only 43 percent of the population is unequivocal with respect to their preferred party—the Likud or the Alignment. Most others either float between the two major parties (Likud and Alignment), between one of those and smaller satellite parties, or do not provide any indication whatsoever on their intended vote.

Table 1 also shows that Israeli pollsters approach this problem with varying degrees of conservatism. The most conservative approach is IASR's approach, which strictly avoids any conclusive classification of floating voters. Instead, it reports the raw percentages obtained in the field, for all possible combinations of voting alternatives considered by subjects. SRC, whose director regularly provides the Israeli TV election night projections, takes a less conservative approach. Although results are still reported in percentages, it seems that all floating votes, except for 15 percent hard to classify subjects, were allocated to one of the party blocks. This classification attempt was greatly simplified by collapsing the fractionalized party system into four major blocks: Likud, Alignment, religious parties, and others.

A much more daring approach is taken by Dahaf and MERI, which are the largest commercial polling organizations in the country. In addition to classifying floaters, both report projections in Knesset seat numbers, present a more detailed list of political parties, and a smaller "other" waste basket category. This indicates elaborate classification schemes on their part allowing relatively finer classification decisions.

However, with increased pressures on the part of the mass media to provide more highly processed and more exciting data to the public, we notice an interesting shift in some of these reporting styles. Subsequent IASR reports in 1981 adopt SRC reporting style and present results

Table 2. Final Projections of Election Results by Israeli Pollsters in 1981 and 1984, and Mean Deviations from Actual Outcomes

	1981				*1984*					
Party	*MERI 6/24–25*	*SRC 6/24–28*	*Dahaf 6/29*	*Election Results 6/30*	*LP*[a] *7/11*	*AP*[b] *7/15*	*Dahaf 7/18*	*MERI 7/19*	*SRC 7/19*	*Election Results 7/23*
Likud	42	43(36%)[c]	43	48	40	36	37	40	41	41
Alignment	42	42(35%)	43	47	45	52	49	46	48	44
L + A	84	85	86	95	85	88	86	86	89	85
L − A	0	1	0	1	−5	−16	−12	−6	−7	−3
Other	36	35(29%)	34	25	35	32	34	34	31	35
A and L overall deviation			5.03					3.91		
L + A deviation			10.03					2.32		
L − A deviation			.82					7.47		

[a] Likud pollster.

[b] Alignment pollster.

[c] The percentages originally reported by the pollsters were translated by this author into seat numbers ($N = 120$) for better comparability and for the purpose of calculating standard deviations.

after a more conclusive classification of floaters. SRC in turn shifted from exclusive percentage-based predictions toward predicting number of seats at the end of the 1984 campaign.

I turn now from the styles of reporting polls which reflect pollsters' refinement efforts to the major issue in this article—accuracy. I define accuracy as the degree of congruence between pollsters' final preelection predictions and actual results. However, for polls conducted earlier in the campaign interpollster consistency may serve as another accuracy measure.

Table 2 relates to the first meaning of accuracy. It presents final predictions of three pollsters in 1981 and five pollsters in 1984, as well as actual election outcomes. The results were collapsed into three categories (Likud, Alignment, and All Others). Two additional rows present the major parties' total number of seats combined (L + A) and the seat difference between them (L − A). The last rows present several accuracy measures. The overall error row presents mean deviations of pollsters' predictions of each of the major parties—from actual results.[6] The L + A error row presents deviations across pollsters of the combined number of seats predicted for the major parties from actual results.[7] Similarly the L − A error row presents mean error predictions of seat difference between the two major parties.[8]

---

[6] The computational formula was:

$$\mathrm{ME} = \sqrt{\sum \frac{(X_{ij} - A_i)^2}{n}}$$

where ME = mean error
i = 1,2 parties
j = pollsters
X_{ij} = last election prediction of pollster j for party i
A_i = actual results for party i
n = total number of predictions.

The "all other" category was excluded in the calculation for convenience reasons since it complements the major parties' number of seats to 120—the number of seats in the Israeli Knesset.

[7] The computational formula is similar to the one presented in footnote 6:

$$\mathrm{MES} = \sqrt{\sum \frac{(X_j - a)^2}{n}}$$

where MES = mean error sum
j = pollsters
n = total number of predictions
X_j = last election prediction of pollster j for (L + A)
a = actual $(L + A)$ result.

[8] The computational formula is the same as in footnote 7, replacing MES with MED (mean error difference) where X_j is the last election prediction of pollster j for $(L - A)$ and a is the actual $(L - A)$ result.

Examining pollsters' overall error in 1981 and 1984, we see that although the 1981 error is larger, the two are not very different. In 1981 MERI, SRC and Dahaf deviated by an average of about five seats in their predictions for the two parties. In 1984 the Likud pollster (LP), the Alignment pollster (AP), Dahaf, MERI, and SRC deviated by an average of about four seats. However, it is evident that the source of these errors is different in the two elections. The major source of error in 1981 lies in inaccurate predictions of the combined number of seats of the two major parties (mean deviation of 10.03), while the difference between the two major parties was accurately predicted (.82 mean deviation). In 1984 the picture is reversed. Errors in the predicted sum number of seats are small (2.32), while most errors lie in the predicted difference between the major parties (7.47). What accounts for these differences? The answer lies in different misclassification patterns stemming from differences in the nature of the floating vote in 1981 and 1984.

From the IASR data in Table 1, it is clear that pollsters' major dilemma in 1981 was to classify the sizeable percentage of voters floating among each major party and its satellites. Many fewer voters floated between the two major parties themselves (Diskin, 1981), accounting for a negligible 3 percent in the IASR data. In light of the close run between the major parties, it seems that there was no justification to apply different classification rules for the Likud and the Alignment floaters. The sizeable error in predicting the total number of seats for the major parties indicates that pollsters consistently erred in allocating more floaters to the smaller rather than the major parties.[9]

The 1984 election differed from 1981 in two respects. First, the competition was not as close as in the previous election. Secondly, many more voters floated between the two major parties. A national poll carried out by the end of June 1984[10] showed that floaters between each major party and its satellites account for 9 percent of the total vote compared to about 18 percent in 1981. However, more than 10 percent of the total vote floated between the Likud and the Alignment, compared to 3 percent in 1981. Although this last percentage seems to have dropped as election came closer, the overall number of floaters between the two major parties was considerably greater than in 1981. In 1984, then, more classification errors occurred between the two major parties and fewer within each party block than in the previous election

---

[9] Some pollsters argued that the polls' predictions themselves showing a close run motivated floaters to lean toward the major parties in their final decision. However, since a strong theory of floaters' classification should be able to account for this possibility as well, we may still consider it a misclassification error.

[10] Dahaf poll.

Table 3. Israeli Polls Results and Interpollster Consistency Measures in the Last Month Before the Election: 1981 and 1984

	Pollster Data Collection: 1981											
	Dahaf 6/1–2	*MERI 6/2–7*	*SRC*[a] *6/4–11*	*IASR*[b] *Beg. June*	*MERI 6/15–18*	*Dahaf 6/16–17*	*IASR 6/14–18*	*Dahaf*[c] *6/26*	*MERI 6/24–25*	*SRC 6/24–28*	*Dahaf 6/29*	*Elect. Results 6/30*
Likud	42	46	45(38%)	19%	49	45	47(39%)	40	42	43(36%)	43	48
Alignment	43	40	40(33%)	24%	37	38	38(32%)	43	42	42(35%)	43	47
L + A	85	86	85		86	83	85	83	84	85	86	95
L − A	−1	6	5		12	7	9	−3	0	1	0	1
Other	35	34	35(29%)	57%	34	37	35	37	36	35(29%)	34	25
S.d. (L + A)			.48			1.25		.5		.5		
S.d. (L − A)			3.09			2.06		1.5		.5		

	Pollster Data Collection: 1984												
	Dahaf 7/1–3	*MERI 7/1–4*	*LP 7/1–4*	*AP 7/8*	*SRC 7/8–10*	*Dahaf 7/11*	*MERI 7/11–12*	*LP 7/11*	*AP 7/15*	*Dahaf 7/18*	*MERI 7/19*	*SRC 7/19*	*Elect. Results 7/23*
Likud	38	37	41	37	35(29.5%)	41	34	40	36	37	40	41	41
Alignment	53	47	48	53	47(39.5%)	50	48	45	52	49	46	48	44
L + A	91	84	89	90	82	91	82	85	88	86	86	89	85
L − A	−15	−10	−7	−16	−12	−9	−14	−5	−16	−12	−6	−7	−3
Other	29	36	31	30	38	29	38	35	32	34	34	31	35
S.d. (L + A)		2.69					3.49			1.41			
S.d. (L − A)		3.67					3.92			2.62			

[a] The percentages originally reported by the pollsters were translated by the author into seat numbers ($N = 120$) for better comparability and for the purpose of calculating standard deviations.

[b] The IASR poll is not comparable to other polls in its group as floating voters remained unclassified. Thus, it was not taken into account in calculating standard deviations.

[c] Another Dahaf poll dated 6/23–24 was excluded from this group since it did not account for the TV debate between the major party leaders as well as a personality change in the alignment leadership—occurring shortly afterwards. The remaining polls in the group accounted in part or fully for these events. The results of this Dahaf poll are very close to the previous Dahaf poll from the 6/16–17.

year. This, of course, explains at least partly the greater accuracy in the sum number of seats for the major parties compared to their projected seat difference.

Table 3 presents 23 polls of major Israeli pollsters published in the mass media in the last month before the elections in 1981 and 1984. The published results are grouped according to polls' date of execution,[11] and collapsed as in Table 2.

If we are willing to accept final results as an accuracy standard for earlier polls as well[12] we can see the same error pattern occurring all along. When we examine the (L + A) and (L − A) rows in Table 3 it is clear that all 1981 polls consistently underestimated the major parties' combined number of seats, but quite accurately predicted the difference between them. In 1984, however, all polls overestimated the difference between the two major parties (L − A), whereas the combined number of seats (L + A) predictions were close to the final results all along.

---

[11] Eight additional polls carried out during May 1981 and June 1984—two months before election date—are not presented here for space reasons. However, they are taken into account in subsequent analysis. In classifying the various polls into groups, we made sure that major political events with a potential effect on public opinion do not separate polls classified in the same group.

[12] This is a problematic assumption suggesting no overtime effect on pollsters' overall prediction accuracy. Testing for this possibility we found that a one-way analysis of variance failed to show significant accuracy differences when contrasting latest predictions with all other predictions. In other words, later clusters did not show significantly greater overall accuracy compared to earlier ones, neither in 1981 nor in 1984.

Turning to the second definition of accuracy, we examined interpollster predictions' consistency within each of the 1981 and 1984 polls clusters. This was done by simply calculating for each cluster the standard deviation from the mean number of seats predicted for the two major parties combined (L + A) and from the mean seat difference predictions (L − A). These deviations are presented in rows six and seven of Table 3.

In contrast to the previous accuracy measure, we do not observe here any interaction of the sum and difference deviations by election year. Instead, standard deviations of (L − A) in the last month before the election are consistently greater than (L + A) deviations, in both years.[13] I suggest that the difference in results evident in the two accuracy measures stems from their sensitivity to different systemic error sources. While deviations from actual results as an accuracy measure reflect floating vote allocation problems, interpollster consistency mainly taps pollster–sponsor dependencies as a source of variability. These dependencies are best reflected in the (L − A) measure, since both politicians and mass media consider the seat difference to be a highly effective cue on campaign momentum in the eyes of the public. Biasing pressures will thus affect the seat difference predictions more than other possible measures, causing greater variability among pollsters on it.

This can be easily seen when predicted seat differences between the major parties are examined in greater detail. For example, Dahaf, working for the Alignment, predicted at the beginning of June 1981 a one-seat difference in favor of its client, while MERI, serving the Likud, predicted its client to lead by a margin of six seats. In the next polling wave, MERI projected an increasing 12-seat gap in favor of the Likud, while Dahaf admitted only a 7-seat difference. By the end of the month competition became close again and was reflected in the 42-seat tie projected by MERI. Dahaf in turn predicted a 3-seat lead for the Alignment. Dahaf's final prediction on June 29 converged to a 43-seat tie while the independent SRC predicted the Likud to lead by one seat. Clearly, then, the systematically biased Dahaf and MERI predictions in favor of their respective sponsors constitute the major source of interpollster variability within time clusters.

The 1984 pattern is no different in this respect. This time we compare the Alignment and Likud free-lance pollsters (AP and LP, respectively). At the beginning of July the Alignment pollster projected a 16-

---

[13] Earlier polls, however—the May polls in 1981 and the June polls in 1984—are not consistent with this observation.

Table 4. Degree of Dependency by Absolute Deviations of (L − A) Predictions from Clusters' Mean for 29 Israeli Polls

Interpollster Accuracy	*Dependency*		
	Low	*Medium*	*High*
0.2 –2.2	71	50	20
2.21–4.2	29	40	20
4.21–6.2	0	10	60
	100%	100%	100%
	(14)	(10)	(5)

$\chi^2 = 12.1$(4 d.f.), $p < .05$ Gamma = .60

seat difference in favor of its client, compared to only a 7-seat difference projected by the Likud's pollster. By mid July the Alignment pollster still stuck to this 16-seat difference while the Likud pollster moved down even further to a narrow 5-seat difference. These mid July projections were also the final AP and LP official predictions released to the press. Other pollsters' final predictions of seat differences all fell between these extremes.

The pattern is very clear—pollsters who were structurally dependent on one of the major political parties favored their sponsor both in 1981 and 1984 election campaigns. With no exception, this pattern also characterizes eight additional polls carried out one month earlier in both campaign years (these polls are not presented in the table for space reasons). In all cases involving partisan relationships, pollsters tended to project results more favorable to their clients or less favorable to their clients' opponents. However, for a more complete support of our hypothesis it is possible to restate the findings of Table 3 somewhat differently, to show the existence of an ordinal relationship between the degree of dependency and accuracy. This is presented in Table 4.

In 1984, the Likud pollster (LP) and the Alignment pollster (AP) worked under highly dependent conditions, operating from sponsors' facilities and working exclusively for their political sponsor. Medium dependency conditions characterized Dahaf's and MERI's environment in 1981. Both operated from their own long established facilities and were commissioned by media organizations in addition to their major political sponsors. SRC's June poll in 1984 (not presented in Table 3) fell into the same category.[14] Totally independent in these respects were SRC in 1981 and Dahaf and MERI in 1984. They did not work for any of the major political parties in the last two months before the election. SRC's July polls in 1984 fit into this category as well.

---

[14] SRC's June poll was carried out still under contract with the Alignment. Soon afterwards SRC stopped working for the Alignment, which retained and enlarged the in-house polling enterprise headed by a university professor on leave.

Table 4 shows the percentage of high, medium, and low accuracy polls in each of the three dependency categories. It is based on 29 polls including those presented in Table 3 and those conducted one month earlier in both election years.[15] Accuracy of each specific poll was measured as its (L − A) absolute deviation from its group's (L − A) mean. The smaller the distance between the cluster's mean and the specific poll's prediction, the greater the accuracy. It is clear that the greater the pollster's dependency, the less accurate are his seat difference predictions. Only one out of the five polls conducted by highly dependent pollsters reported parties' seat difference accurately. The results of Table 3 do not leave any doubt as to the direction of bias. Fifty percent of the medium-dependent pollsters' polls were accurate in this sense, while more than 71 percent of independent pollsters' polls were highly accurate.

Discussion

Our findings suggest that processes affecting accuracy in partisan-sponsored polls are more complex than the Hennessy and Hennessy (1961) contention that pollsters wish to please their partisan employers. The fact that pollsters' tendency to favor their sponsors varies monotonically with the strength of their structural dependencies points to the role of structural factors affecting accuracy. For example, it is likely that prolonged hours of physical proximity to sponsors' campaign facilities and close social ties with campaign officials expose pollsters to high levels of optimistic expectations as to the campaign outcome. Thus, the closer the relationships, the stronger the biasing social pressures directed toward pollsters. Under such conditions it is perhaps easier to interpret vague data in a favorable rather than an unfavorable direction to surrounding expectations. Although not directly tested in this study, we suspect that pollsters' data refinement efforts in a low transparency system, based partly on subjective judgment and experience, are not rigid enough to withstand environmentally biasing pressures.

The role of pollsters' data refinement efforts as a source of error becomes more evident when final election predictions are examined. The different nature of pollsters' prediction errors across election years can be partly attributed to classification errors of the sizeable floating vote in the system. Thus, attempting to provide greater detail and refinement based on relatively weak theories in a low transparency system, researchers expose themselves to greater risk of error. This raises an important normative question with respect to pollsters' role in election times. Clearly, in the Israeli context pollsters were driven both by competition and by a "horserace" attitude on the part of the press to provide highly processed projections that are also more exciting. Their failure resulted in a decline in status and prestige of the entire polling industry. An alternative course that seems more promising, although less glossy, is to focus on descriptive and explanatory aspects of the data as long as existing behavioral theories do not provide a more sound base for accurate predictions.

---

[15] Two IASR polls were not included in the analysis as they did not attempt to classify the floating votes and are thus not comparable to the rest.

References

Arian, A.
1981 "Elections 1981: competitiveness and polarization." Jerusalem Quarterly 21:3–27.
Broh, A. C.
1980 "Horse race journalism: reporting the polls in the 1976 presidential election." Public Opinion Quarterly 44:514–29.
Cantril, A. H.
1976 "The press and the pollster." Annals of the Academy of Political and Social Science 427:45–52.
Crespi, I.
1980 "Polls as journalism." Public Opinion Quarterly 44:462–76.
Diskin, A.
1981 "The 1981 election: public opinion polls." Jerusalem Quarterly 22:99–104.
Diskin, A., and D. S. Felsenthal
1981 "Do they lie?" International Political Science Review 2:407–22.
Field, M. D.
1971 "Towards responsibility in reporting opinion surveys, the researchers' view." Public Opinion Quarterly 35:342–46.
Hennessy, B. C., and E. R. Hennessy
1961 "The prediction of close elections: comments on some 1960 polls." Public Opinion Quarterly 25:405–11.
Meyer P.
1973 Precision Journalism, A Reporter's Introduction to Social Science Methods. Bloomington: Indiana University Press.
Miller, M. M., and R. Hurd
1982 "Conformity to AAPOR standards in newspaper reporting of public opinion polls." Public Opinion Quarterly 46:243–49.
Noelle-Neuman, E.
1980 "The public opinion research correspondent." Public Opinion Quarterly 44:585–97.
Perry, P. K.
1962 "Correspondence." Public Opinion Quarterly 26:133–35.
Tuchman, G.
1978 Making News: A study in the Construction of Reality. New York: The Free Press.
Weimann, G.
1982 "Press Coverage of pre-election polls in Israel." State, Government and International Relations 19–20:132–43 (Hebrew).

Mass Communication and Political Socialization: Specifying the Effects

GINA M. GARRAMONE AND CHARLES K. ATKIN

STUDIES examining the role of mass communication in young people's learning about politics indicate that the media contribute to political socialization, but the research does little to unravel the complexities of the process. By treating the media exposure patterns, the criterion variables, and the audience as unidimensional concepts, these studies fail to tease out the more specific effects that might exist for *certain* media, *certain* variables, or *certain* audience categories. The present study attempts to address these shortcomings by comparing the socialization effects of exposure to four media sources of political information—television news, radio news, newspaper news, and newsmagazines—on several specific types of political knowledge and behavior. The study also explores differences in effects for certain subgroups of young people.

Abstract A survey investigation of teenagers examined differential socialization effects of four news media on four types of political knowledge and behavior among several categories of youth on the basis of abilities and predispositions. Regression analyses of questionnaire responses from 280 seventh and tenth grade students show that TV news exposure is the strongest predictor variable. The broadcast media have a greater impact on current events knowledge than fundamental political knowledge, while a slight reversal occurs for the print media. Newspaper reading has the greatest effect on participation. There are predominantly uniform relationships across subgroups, providing little evidence that the news media produce knowledge or behavior gaps.

Gina Garramone is Assistant Professor of Advertising and Charles Atkin is Professor of Communication and Telecommunication, both at Michigan State University. The research was supported by the National Association of Broadcasters.

Public Opinion Quarterly Vol. 50:76–86 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-76/$2.50

Four Types of Media

Although entertainment and persuasive content in the mass media contain some politically relevant messages, news items provide the bulk of political information. A number of studies demonstrate widespread consumption of hard news by young people (Gollin and Anderson, 1980; Drew and Reeves, 1980a; Atkin, 1978). The inputs from news media usage do not constitute a unidimensional phenomenon, however. The four news media differ in terms of inherent *channel mode* attributes, conventional *style* of news presentation, and nature of *content*.

The television medium can present visual images, while the other media are largely restricted to verbal modes. The print media allow considerable receiver control of information selection and processing, requiring some degree of motivated and focused attention. The sender-controlled broadcast transmissions require minimum audience effort (Chaffee and Tims, 1982).

Stylistically, the print news media tend to feature greater complexity with respect to technical readability and substantive intricacy; broadcast news is presented in a simplified and concrete style that facilitates comprehension. In addition, the dramatic, personalized, and entertaining presentational style of television news is highly engaging relative to the other media.

While the topical content differs minimally across media, newspapers convey a larger quantity of messages, and print news items are characterized by greater depth coverage, with fuller detail, broader perspective, and richer background information compared to the ahistorical and abbreviated broadcast content.

Each of these dimensions has implications for differential impact on adolescents at the cognitive and behavioral levels.

Political Criterion Variables

KNOWLEDGE

Past studies have found a relationship between media exposure and young persons' current events knowledge (Chaffee et al., 1970; Atkin and Gantz, 1978) and more fundamental knowledge about the political scene (Conway et al., 1975; Gollin and Anderson, 1980). *Current events knowledge* comprises knowledge of the discrete, time-bound events and associated individuals featured in the news media. In the 1981 period of this study, for example, IRA hunger-striker Bobby Sands died. Such concrete, specific events and the names of the individuals associated with them typically receive heavy media coverage. These events are not independent of the political history and extant political milieu in the respective country, of course. The historical conflict in Northern Ireland, including the religious and political orientations of the combatants, provides a certain perspective on the hunger strikers. Such *fundamental knowledge*, of political philosophies, structures, and historically significant developments, constitutes a more basic understanding underlying daily political events.

Most studies have limited their analyses to only one type of knowledge and only one type of medium, precluding comparisons of differential effects. On the basis of the mode, style, and content differences between broadcast and print media, however, we may predict differential effects likely to emerge. Television as a news medium, for example, compels contact to each story, facilitating exposure to a variety of topics. Compared to print, broadcast news stories tend to be briefer, simpler, easier to process, and more engaging in presentational style. All of these factors suggest that the broadcast media, and television in particular, should be especially effectual for young audiences, who tend to be visually oriented and less cognitively sophisticated than adults. We predict, therefore, that broadcast news exposure will be more strongly related to overall political knowledge than will print news exposure.

Chaffee and Tims (1982) noted that exposure to the superficial broadcast media may be particularly effective for learning such "itemized" cognitions as the names of political actors or parties. Regular use of print media, on the other hand, with their depth of coverage, substantive complexity, and self-paced processing, may be more likely to enable the youngster to acquire and maintain complex cognitive associations between the various elements of the political scene. We predict, therefore, that exposure to broadcast news will contribute more to knowledge regarding current events than to fundamental knowledge, while print news exposure will have greater impact upon fundamental than on current events knowledge.

BEHAVIOR

Although youngsters cannot participate formally in the political system, they can practice interpersonal discussion behaviors and form future participatory dispositions. The news media provide raw information that can be conversationally shared with others (Atkin and Gantz, 1978; Drew and Reeves, 1980a). By providing youngsters with information regarding others' political behavior, mass media exposure also may contribute to the development of long-range predispositions regarding such anticipated political activities as voting and campaign work (Sheinkopf, 1973). The mode, style, and content characteristics of the broadcast and print classes of media may also determine their differential impact on the two types of political behavior. The broadcast media, with their dramatic and entertaining presentation style, should have greater impact on interpersonal discussion than on the more complex anticipated participation. The more involving and educative print media, with their detail and broader perspective, should be more strongly related to anticipated participation than to discussion.

Categories of Young People

The various capabilities and orientations of the young media audience will influence media effects. For example, teenagers possessing greater cognitive sophistication are predicted to acquire complex political information more broadly and deeply than do younger children (Atkin and Gantz, 1978; Drew and Reeves, 1980b). The impact of cognitive sophistication on learning from the media may also depend upon the type of medium under consideration; younger and older students should learn equally well from the *broadcast* media, but older students should outperform younger students in learning from the more complex *print* media. The organismic variable of age is most often used as an indicator of cognitive sophistication. A child's scholarly orientation (e.g., grades in school, intention to attend college) is also likely related.

Interest in politics is another significant intervening variable. Youngsters with greater interest should more readily respond to news coverage. The impact of interest, however, may depend upon the class of media under consideration. Interest may have less impact on learning from the less involving broadcast media than from print media.

Finally, antecedent variables which may shape the political learning process are assessed. Research indicates that boys are more attuned to political news and more involved politically than are girls, and that socioeconomic status is slightly related to these variables (Atkin, 1981).

In summary, it is predicted that broadcast news exposure will be more strongly related to overall political knowledge than will print news exposure. It is also proposed that broadcast news exposure will be a better predictor of current events knowledge than of fundamental knowledge, but that print news exposure will be a better predictor of fundamental than of current events knowledge. Similarly, it is predicted that broadcast exposure will be more closely related to interpersonal discussion than to anticipated participation, but that print exposure will be more closely related to anticipated participation than to discussion. In addition, the relationships between media exposure and the criterion variables are expected to vary for subgroups of the young audience at various levels of age, scholarly orientation, and political interest.

Method

Questionnaires were administered to public school students in two mid-sized midwestern cities with typical socioeconomic characteristics. Students were drawn from the seventh grade ($N = 149$) and tenth grade ($N = 131$) classrooms at each site. The questionnaire contained items measuring the four types of media exposure, political knowledge and behavior, age, scholarly orientation, political interest, gender, and socioeconomic status. The variables were operationalized in the following manner:

Television News Exposure. The index of exposure to television news summed items dealing with how often national network newscasts were viewed, how many days had been watched the previous week, whether the news was viewed the day before the survey, the number of days viewed in an average week, and the proportion of the program that was usually watched; in addition, the frequency of viewing morning news-variety programs was included.[1]

Radio News Exposure. The radio news listening index summed the estimated number of radio news shows heard on an average day, on the previous day, and the amount of attention paid to radio news stories.

Newspaper News Exposure. This index summed respondents' answers to how often they read the front page of the newspaper and how many hard news articles they read per day.

Newsmagazine Exposure. This was measured by frequency of reading the weekly periodicals such as *Time* and *Newsweek.*

Some analyses call for a comparison of the two classes of broadcast and print media. For these analyses, television news exposure and radio news exposure are summed to form a *broadcast exposure* index. Newspaper news exposure and newsmagazine exposure are summed to form a *print exposure* index.

Political Knowledge. An index for *overall* political knowledge was constructed by summing together 46 questionnaire items, most of the "multiple choice" variety. The typical item posed a question about events, developments, or people in the political sphere, and listed three choices and the fourth alternative of "??" which students were encouraged to use if they were unsure.[2] This overall knowledge index is used in the contingent correlation analysis.

---

[1] The means, standard deviations, and Cronbach's alpha reliability coefficients for the indices are as follows: television news exposure, 17.19, 8.35, .78; radio news exposure, 6.31, 2.82, .71; newspaper news exposure, 4.72, 1.69, .62; newsmagazine exposure, 1.62, .67, estimated at .50; broadcast exposure, 23.50, 9.48, .78; print exposure, 6.34, 1.97, .63; overall political knowledge, 70.47, 8.64, .89; fundamental knowledge, 10.93, 2.08, .68; current events knowledge, 12.88, 2.34, .69; interpersonal discussion, 16.69, 5.04, .88; and participation, 6.32, 2.04, .51.

From the 46 knowledge items, the two more specific knowledge indices were constructed. The *fundamental* knowledge index was created by summing together eight items concerning historical facts or trends relevant to more current events.[3] To make a *current events* knowledge index comparable to the fundamental knowledge index, eight current news items were selected to match as closely as possible those measured in the fundamental knowledge index.[4]

Political Behavior. The *interpersonal discussion* index is composed of eight items dealing with interpersonal communication with both family and friends about four news topics.[5] The *participation* index measures anticipated election campaign activities, inclination to run for office, and plans to vote in elections.

Contingent Variables. Three variables were measured for contingent analyses: age, scholarly orientation, and political interest. *Age* is indicated by the student's grade in school. *Scholarly orientation* is measured by an index summing self-reported performance in school classes and intention to attend college. The *political interest* index measures how interested the student is in several news topics.[6] For use in the contingent analyses, each of these variables was divided at the median.

---

[2] Sample knowledge items include: "President Reagan recently made a trip to give a speech, which was the first time he left Washington since he was wounded. Where did he travel to?" "The leader of Japan visited Washington to talk with President Reagan. What did they talk about?" "What is the name of the man who is Budget Director for the U.S.?" "A few months ago, the 52 hostages came back to the United States. In what country were the hostages held prisoner?" "How many hunger strikers have died in Northern Ireland prisons?" "There have been debates about whether the U.S. should send military aid to a Central American country. Which country?" "Which country just elected a new president?" "What does Ronald Reagan want to sell to Saudi Arabia?"

[3] Sample items are: "Poland is run by a Communist government," "The conflict in Northern Ireland is caused by problems between two religious groups."

[4] Sample items are: "What are the people of Poland worried about this year?" "A man in Ireland named Bobby Sands died several weeks ago. How did he die?"

[5] The four topics are: President Reagan, things that happen in other countries (Poland, Russia, Israel, Iran, El Salvador are listed as examples), government and things that happen in Washington, and current problems in the news (crime, inflation, nuclear power, gun control are listed as examples).

[6] The topics included "things that happen in other countries," "things that President Reagan does," "government and things happening in Washington," and politics in general.

Table 1. Zero-order and Corrected Correlations Between Media Exposure and the Criterion Variables.[a]

	Knowledge			*Behavior*	
	Overall	*Current Events*	*Fundamental*	*Interpersonal Discussion*	*Anticipated Participation*
Broadcast	.43 (.52)	.39 (.53)	.26 (.35)[b]	.45 (.54)	.29 (.46)[c]
Print	.32 (.43)[d]	.24 (.36)	.24 (.36)	.39 (.52)	.37 (.65)[e]

[a] The correlations corrected for attenuation (Nunnally, 1967) are represented within parentheses. Corrections are made based on the internal consistency estimates (Cronbach's alpha) for each index.

[b] For Broadcast exposure, the difference between Current Events versus Fundamental knowledge attenuation-corrected correlations is significant at the 5 percent level.

[c] For Broadcast exposure, the difference between Interpersonal Discussion versus Anticipated Participation attenuation-corrected correlations is significant at the 5 percent level.

[d] For Overall Knowledge, the difference between Broadcast versus Print exposure attenuation-corrected correlations is significant at the 5 percent level.

[e] For Print exposure, the difference between Interpersonal Discussion versus Anticipated Participation attenuation-corrected correlations is significant at the 5 percent level.

Background Variables. In addition to the contingent variables, the student's gender and family socioeconomic status[7] were measured for control purposes. In some analyses, age, scholarly orientation, and political interest are also treated as background predictor variables.

ANALYSIS

Correlation and regression analysis are used to explore the relationship between media exposure and each of the criterion variables. Contingent correlations between media exposure and the criterion variables at high versus low levels of age (grade), scholarly orientation, and political interest also are computed.

Results

The results are described in the following order: (1) the relationship between media exposure and political knowledge, (2) the relationship between exposure and political behavior, (3) the contingent analyses.

POLITICAL KNOWLEDGE

Zero-order correlations were computed between broadcast and print classes of media exposure and overall political knowledge, current events knowledge, and fundamental knowledge. Because of the different reliabilities of the indices, the correlations are corrected for attenu-

---

[7] Socioeconomic status is indexed by parental occupation, social class of neighborhood, and family member college education.

Table 2. Regression of Criterion Variables on Demographic and Media Variables

	Knowledge			*Behavior*	
	Overall	*Current Events*	*Fundamental*	*Interpersonal Discussion*	*Anticipated Participation*
SES	.08[a]	.06	.08	−.04	.06
Gender[b]	.25	.26	.13	.06	−.06
Grade[c]	.27	.28	.27	.03	−.16
Scholarly orientation	.39	.28	.31	.38	.33
(R^2 for step)	(.33)	(.25)	(.23)	(.14)	(.14)
Television	.27[d]	.27	.16	.25	.08
Radio	.07	.10	.02	.14	.08
Newspaper	.04	.02	.05	.19	.22
Newsmagazine	.09	.04	.09	.04	.07
(Incremental R^2 for step)	(.11)	(.09)	(.04)	(.17)	(.09)

[a] Coefficients in the first four rows are standardized beta weights controlling for other demographic variables; $\beta > .11$ significant at 5 percent level.

[b] Gender is coded female-male.

[c] Grade in school is coded 7th-10th.

[d] Coefficients in the bottom rows are standardized beta weights controlling for demographic variables and other exposure variables; $\beta > .11$ significant at 5 percent level.

ation (Nunnally, 1967) using the internal consistency estimates. Both sets of correlations are presented in Table 1. The description of results and tests of significance pertain to the attenuation-corrected correlations. As expected, broadcast news exposure is more strongly correlated with overall political knowledge than is print news exposure (t (3,277) = 2.78, $p < .05$). While broadcast exposure also is strongly correlated with current events knowledge, it is significantly less strongly correlated with fundamental knowledge (t (3,277) = 4.03, $p < .05$). For print news exposure, the correlations with current events and fundamental knowledge are identical.

Hierarchical regressions were computed to provide a more detailed analysis of the relationship between the media exposure and knowledge variables, with background variables controlled. Entered in the first step of the regressions were socioeconomic status, gender, grade in school, and scholarly orientation (Table 2). The four media exposure indices were added as a block for the second step.

Three of the four background variables are significantly related to overall political knowledge. Males, tenth graders, and students with a strong scholarly orientation are more knowledgeable politically. Of the four media exposure variables, the only significant predictor is television news viewing, with a standardized beta coefficient (β) of .27. The next strongest predictor is newsmagazine reading.

Corresponding regression analyses were computed for current events and fundamental knowledge. As a block, the four media variables explain twice as much variance in current events knowledge ($IR^2 = .09$) as in fundamental knowledge ($IR^2 = .04$). Television news viewing is the best predictor of both current events and fundamental knowledge, but the relationship is much stronger for the former. Radio news listening is also a stronger predictor of currents events knowledge than of fundamental knowledge. Both of the print media indices, however, are slightly better predictors of fundamental knowledge than of current events knowledge. Addition of the political interest variable into the knowledge regression equations has little impact on the media coefficients.

POLITICAL BEHAVIOR

Zero-order and attenuation-corrected correlations were computed between media exposure and interpersonal discussion and anticipated participation (Table 1). Again, the description of results and tests of significance pertain to the attenuation-corrected correlations. The predicted pattern of differences between the discussion and participation correlations emerges: broadcast news exposure is more strongly related to interpersonal discussion than to anticipated participation ($t\ (3,277) = 1.66$, $p < .05$), and print news exposure is more strongly related to anticipated participation than to discussion ($t\ (3,277) = 3.07$, $p < .05$).

Hierarchical regressions were computed to assess the relationship between media exposure and the political behavior indices, with background variables entered first (Table 2). Students with a strong scholarly orientation are significantly more likely both to engage in interpersonal discussion regarding politics and to anticipate future political activity. Seventh graders are more likely to anticipate participation than are tenth graders.

The four media variables make a substantial contribution to predicting interpersonal discussion ($IR^2 = .17$). Television news viewing is once again the strongest predictor ($\beta = .25$), with newspaper news reading and radio news listening also significant predictors. For anticipated participation, newspaper news reading is most closely related ($\beta = .22$), with the remaining three indices failing to reach significance. Again, addition of the interest variable into the regression equations has a negligible impact on the media coefficients.

CONTINGENT ANALYSES

Contingent correlations were computed between the broadcast and print exposure and criterion variables at high versus low levels of grade, scholarly orientation, and political interest. The expectation that the contingent variables would significantly mediate the relationship between print media exposure and the criterion variables was not supported. Only one such contingent relationship emerges, and in the direction opposite to that predicted. The correlation between print media exposure and anticipated participation is stronger for students at a low level of scholarship than for those at a high level. This contingent relationship is also evident for broadcast media exposure.

Discussion

The present study sought to examine the impact of four types of media exposure—television news, radio news, newspaper news, and newsmagazines—on two types of political knowledge and two forms of political behavior. In addition, the relationships between media exposure and the criterion variables were examined for subgroups of young persons at various levels of age, scholarly orientation, and political interest.

The results of both correlation and regression analyses support the expectation that broadcast news exposure is more strongly related to overall political knowledge than is print news exposure. Also as predicted, broadcast exposure is a better predictor of current events knowledge than of fundamental knowledge. Contrary to expectations, however, print news exposure predicts the two types of knowledge equally well. The predictions regarding the behavioral variables were supported by the data. Broadcast news exposure is more closely related to interpersonal discussion than to anticipated participation, and print news exposure is more strongly related to participation than to discussion.

Only the prediction that print news exposure would be more strongly related to fundamental than to current events knowledge failed to be supported by the data. This failure may be attributed to some of the same characteristics of print news media that led to the original prediction. These characteristics, when interpreted in light of youngsters' limited information-processing abilities, become disadvantages rather than advantages of print for youngsters' gaining fundamental political knowledge. For example, the *in-depth coverage* of print news may result in information overload for the less developed mind. Similarly, the *complexity* of print news stories may render them less comprehensible. The advantage of *self-paced processing* of print applies only to motivated people; the less interested youth may simply fail to capitalize on it.

If processing limitations inhibit youngsters' learning from the print media, then variables indicative of level of processing may mediate fundamental knowledge gain from print news media. Thus older students, brighter students, and more interested students might more efficiently extract fundamental knowledge from the broadcast media than the younger, less bright, and less interested students. The contingent relationships for fundamental knowledge (not shown) do not reveal any consistent pattern in support of this argument, however. But age, scholarship, and political interest are only indirect indicators of processing. More direct measurement of processing suggests that "intense" processing facilitates learning from television news (Garramone, 1983), but whether it facilitates learning sufficiently to allow the extraction of the more fundamental political knowledge should be tested.

The overall pattern of findings across subgroups provides little evidence of unequal effects of exposure; the news media do not appear to produce knowledge gaps (or behavior gaps) between those adolescents who are brighter or politically interested and those who are less scholarly or uninterested. If anything, the media serve a compensatory function in some cases by more strongly stimulating the political orientations of exposed students scoring low on these background variables.

References

Atkin, Charles K.
1978 "Broadcast news programming and the child audience." Journal of Broadcasting 22:47–61.
1981 "Communication and political socialization." Pp. 299–328 in Dan Nimmo and Keith Sanders (eds.), Political Communication Handbook. Beverly Hills, CA: Sage.
Atkin, Charles K., and Walter Gantz
1978 "Television news and the child audience." Public Opinion Quarterly 42:183–98.
Chaffee, Steven H., Scott Ward, and Leonard Tipton
1970 "Mass communication and political socialization." Journalism Quarterly 47:647–59.
Chaffee, Steven H., and Albert R. Tims
1982 "The development of news media use in adolescence: implications for political cognitions." In Michael Burgoon (ed.), Communication Yearbook 6. Beverly Hills, CA: Sage.
Conway, M. Margaret, A. Jay Stevens, and Robert Smith
1975 "The relation between media use and children's civic awareness." Journalism Quarterly 52:531–38.
Drew, Dan, and Byron Reeves
1980a "Children and television news." Journalism Quarterly 57:45–54.
1980b "Learning from a television news story." Communication Research 7:121–35.
Garramone, Gina M.
1983 "Television news and political socialization." In Robert Bostrom (ed.), Communication Yearbook 7. Beverly Hills, CA: Sage.
Gollin, Albert, and Thelma Anderson
1980 America's Children and the Mass Media. New York: Newspaper Advertising Bureau.
Nunnally, Jum
1967 Psychometric Theory. New York: McGraw-Hill.
Sheinkopf, Kenneth
1973 "Family communication patterns and anticipatory socialization." Journalism Quarterly 50:24–30.

Not Forbidding Isn't Allowing: The Cognitive Basis of the Forbid-Allow Asymmetry

HANS-J. HIPPLER AND NORBERT SCHWARZ

SURVEY RESEARCH has repeatedly demonstrated that minor changes in the wording of a question can have a major impact on the obtained responses (see Kalton and Schuman, 1980, and Schuman and Presser, 1981 for reviews). A wording effect that has received much attention in more than four decades of research is the forbid-allow contrast identified by Rugg (1941). He reported that Americans were more likely to support freedom of speech when the appropriate question was worded, "Do you think the United States should forbid public speeches against democracy" (yes/no?) rather than, "Do you think the United States should allow public speeches against democracy" (yes/no?). Specifically, only 54 percent of the respondents who were asked

---

Abstract Previous research demonstrated that respondents are more likely to endorse the idea that something should "not be allowed" (or "not be forbidden") than to endorse the idea that it should be "forbidden" (or "allowed"), even though these expressions seem logically equivalent. The hypothesis is advanced that this asymmetry is due to the response behavior of indifferent respondents who neither endorse that something should be forbidden nor that it should be allowed, resulting in higher endorsements of the negative form of both question wordings. Data consistent with that explanation are presented and the cognitive mechanism underlying the response behavior of indifferents is discussed.

Hans-J. Hippler is Project Director, Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim, and Norbert Schwarz is Assistant Professor of Psychology, Universität Heidelberg, Federal Republic of Germany. This research was partially supported by the Zentrum für Umfragen, Methoden und Analysen and grant Schw 278/2 from the Deutsche Forschungsgemeinschaft to Norbert Schwarz and Fritz Strack. The authors gratefully acknowledge the stimulating discussions with Charles F. Cannell, Howard Schuman, Seymour Sudman, and Fritz Strack. Correspondence should be addressed to H. J. Hippler, ZUMA e.V., Postfach 5969, D-6800 Mannheim 1, West Germany.

Public Opinion Quarterly Vol. 50:87–96 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-87/$2.50

the forbid form of the question felt that those speeches should be "forbidden," but 75 percent of the respondents who were asked the allow form of the question felt that they should "not be allowed." Similarly, only 25 percent of the respondents felt that these speeches should be "allowed," but 46 percent responded they should "not be forbidden."

Since this original study, the effect has been successfully replicated over a wide range of issues, and Schuman and Presser (1981:296), in reviewing this work, noted that it is "the largest wording effect we have discovered," producing differences of up to 30 percentage points. Three factors that possibly contribute to the effect have been suggested. First, and most notably, "forbid" and "allow" carry different connotations: "the former sounds harsher and may therefore be more difficult to endorse, whereas the latter in some context might seem to encourage a deviant behavior and therefore may invite opposition" (Schuman and Presser, 1981:280). Second, the terms create different grammatical structures, and finally, the effect may be limited to abstract issues such as speeches against democracy. Unfortunately, none of these explanations received unequivocal support in studies reported by Schuman and Presser, and they conclude that they "are hard put to see any obvious source for the effect" (1981:296).

However, it seems to us that these authors discarded the connotation explanation somewhat prematurely on the basis of a potentially inadequate test. Specifically, Schuman and Presser (1981:281) found large and reliable wording effects for abstract items (e.g., speeches against democracy) but small and unreliable effects for more concrete items (e.g., the showing of X-rated movies, or cigarette advertisement on TV). They concluded from this pattern: "If forbid and not allow (or not forbid and allow) convey somewhat different meanings, the effect of the verbal difference taken alone is not large. For the alternative formulations to have much impact, it would appear that they have to be joined to abstract issues like free speech or communism" (1981:282). Interestingly, Schuman and Presser also reported higher don't know levels for their abstract than for their concrete items, and they concluded that "the higher the don't know level the larger the forbid-allow effect" (p. 283). To the extent that high don't know levels indicate a high level of indifference in the public opinion climate (Noelle-Neumann, 1980), this finding suggests a variant of the connotation explanation that might well account for the data reviewed by Schuman and Presser (1981) in the manner outlined below.

Specifically, we want to suggest that the forbid-allow asymmetry is primarily due to the response behavior of indifferent (or ambivalent) respondents. We assume that respondents who hold a strong attitude toward the issue under study will give consistent responses to both forms of the question. That is, respondents opposed to the issue will endorse the forbid statement as well as the not allow statement, while those in favor of the issue will endorse both the allow and not forbid statements. However, indifferent respondents may respond "no" to *both* questions. Specifically, when asked whether something should be forbidden, indifferents may say "no" because they are not opposed to the issue but only indifferent. However, when asked whether it should be allowed they may respond "no" for the same reason—even though a "no" response seems to imply different opinions depending on which question it is given to. Two factors may contribute to this phenomenon, namely, the connotations of "forbid" and "allow," and a cognitive bias to focus on positive instances of a behavior rather than to consider the implications of its absence.

With regard to the connotations of "forbid" and "allow" it has already been noted that to "forbid" is a rather harsh statement while "allowing" something may be construed as inviting it, rather than tolerating it. That is, to *forbid* something implies stronger opposition than *not to allow* something, because forbidding refers to an active act of opposition against the issue under consideration while not allowing something only refers to abstaining from active support. Similarly, *allowing* something refers to active support while *not forbidding* something refers to abstaining from active opposition.

Thus, respondents who do not hold a strong attitude may be unlikely to endorse *either* the forbid or the allow statement, responding "no" to both question forms. In doing so they seem to endorse the opposite position. However, they may miss this in the interview situation because they focus on the behavior they are asked about. In this regard, research in the area of social cognition demonstrated that individuals pay more attention to positive instances of a behavior than to the absence of a behavior (cf. Nisbett and Ross, 1980 for a review). For example, they feel that what a person does do provides more information about the person than what the person does not do (e.g., Fazio et al., 1982). This "tendency of both animals and humans to exhibit greater difficulty in the processing of nonoccurrences than occurrences" (Fazio et al., 1982:404) is usually referred to as the *feature-positive effect.* Similarly, when asked if something should be done (allowed or forbidden), individuals may focus on the implications of the behavior under consideration, and they may *not* consider the implications of the *absence* of this behavior. Such a bias could contribute to the forbid-allow asymmetries by clouding the idea that not allowing or not forbidding something indeed amounts to forbidding or allowing it, respectively, thus resulting in the feeling that one has only said it shouldn't be allowed without ever commenting on whether it should be forbidden.

Note that this explanation predicts that the asymmetries are primarily due to the behavior of indifferent or ambivalent respondents and should not be obtained for respondents holding a clear pro or con attitude. In line with this hypothesis, the forbid-allow effect was found to be more pronounced the more abstract the issue and the lower the education of the respondents, both factors that may affect the number of indifferent respondents, which is also partially reflected in increasing don't know levels.

In this article we will report evidence that to forbid and to allow are stronger statements than not to allow or not to forbid (Experiment 3), and that the forbid-allow asymmetry is limited to indifferent respondents and is not obtained for respondents holding a strong attitude (Experiment 4). Because our studies were conducted in West Germany, however, we will first demonstrate that the basic forbid-allow effect is cross-culturally replicable (Experiments 1 and 2).

Experiments 1 and 2: The Forbid-Allow Effect in German Samples

To test the cross-cultural stability of the forbid-allow effect, two split-ballot experiments were conducted. The first experiment was carried out in July 1982 as part of the second wave of a small face-to-face panel survey with a random sample of adults (more than 18 years of age) living in an industrial city in West Germany and selected from registries (response rate first wave: 63 percent; second wave: 89 percent). The second experiment was carried out in April 1983 as part of a larger face-to-face survey (quota sample) with 146 German adults. Quotas were based on an intersection of age, sex, and years of education. In the first experiment, 48 of 88 respondents were asked if peep shows should be allowed (yes or no), and 40 were asked if they should be forbidden (yes or no). To provide a stimulus replication, 146 respondents in the second survey were asked if the showing of X-rated movies in public cinemas should be allowed ($N = 72$) or forbidden ($N = 74$). Both peep shows and X-rated movies are legal in West Germany.

Both studies replicated the forbid-allow asymmetry, demonstrating its cross-cultural stability. Specifically, 27.7 percent and 26 percent of the respondents, respectively, endorsed the statement that peep shows and the showing of X-rated movies should be forbidden, while 50 percent and 40 percent, respectively, reported they should not be allowed, χ^2 (1) = 3.5 and 2.9, respectively, $p < .06$ and .10 (three and two subjects, respectively, provided a don't know response). Thus, differences of 23.3 and 14 percentage points were obtained as a function of question wording, using two issues that qualify as concrete rather than abstract in Schuman and Presser's terminology. As in Schuman and Presser's (1981:282) data for X-rated movies, the effects were only marginally reliable. However, a combined analysis of both studies (following a suggestion by Rosenthal, 1978) indicates the statistical reliability of the general forbid-allow asymmetry in both German samples, $z = 2.49$, $p < .01$, one-tailed.

Experiment 3: Forbidding and Allowing Are Stronger Statements Than Not Allowing or Not Forbidding

To test the hypothesis that to allow something or to forbid something are stronger positions to take than not to forbid or not to allow it, ratings of statements containing these expressions were obtained.

METHOD

As part of a self-administered questionnaire 54 adults (more than 18 years of age) of a quota sample (quotas were based on an intersection of age, sex, and years of education) read a sentence stating that "Mr. Mueller feels the legislator should forbid (not allow, allow, not forbid) peep shows," and subsequently rated Mr. Mueller's opinion on an 11-point bipolar scale labeled "extremely in favor of peep shows" vs. "extremely opposed to peep shows." Each respondent read and rated only one statement.

RESULTS

As expected, "allowing" peep shows was rated as representing a more favorable attitude, $M = 2.1$, than "not forbidding" peep shows, $M = 3.1$, $t(50) = 2.14$, $p < .04$. Similarly, "forbidding" peep shows was rated as indicating stronger opposition, $M = 10.5$, than "not allowing" them, $M = 9.2$, $t(50) = 2.69$, $p < .01$, demonstrating that the terms are perceived as differing in extremity even though they may be logically equivalent in their consequences. The differences are small, however, and there is no doubt that "not forbidding" is seen as a statement in favor of the issue while "not allowing" is seen as a statement in opposition to the issue.

Experiment 4:
The Forbid-Allow Effect Is Limited to Indifferent Respondents

Central to the present argument is the assumption that indifferent respondents may respond negatively to both a question asking them whether something should be allowed and a question asking them whether it should be forbidden. Moreover, it is assumed that this response behavior accounts for the higher overall endorsement of the negative form of both questions. To test this hypothesis, respondents in favor of, opposed to, or indifferent toward spreading salt on the roads for safety in winter—a popular topic in the current ecological discussion in West Germany—were asked whether using salt to melt ice and snow should be allowed (yes or no) or forbidden (yes or no), resulting in a 3 × 2 factorial design with attitude (pro, con, indifferent) and question wording (forbid, allow) as between subject factors.

It was expected (a) that more respondents overall would endorse the "not allow" than the "forbid" and the "not forbid" than the "allow" statements, and (b) that this would be primarily due to the responses of indifferent subjects, and would not be obtained in the responses of subjects holding a strong attitude.

METHOD

A random sample (selected from registries) of 720 adults (more than 18 years of age) living in an industrial city in the Federal Republic of Germany participated in a mail survey (response rate 78 percent) concerned with environmental issues. Early in the questionnaire (question number 10) respondents reported their attitudes toward 12 environmental issues, including the use of salt for road safety in winter. Specifically, the question read, "Salt should no longer be used to melt ice and snow on the roads." Respondents were asked to check one of three response alternatives ("agree," "don't agree," "indifferent/don't care").

Later in the questionnaire (question number 34) half of the respondents were asked whether the use of salt to melt ice should be allowed (yes or no), while the remaining respondents were asked whether it should be forbidden (yes or no). The question read: "The use of salt during the winter is dangerous to plants and trees near the road. Do you think that one should generally allow (forbid) the use of salt (yes/no)?" A "don't know" option was not provided with this question.

RESULTS

Table 1 presents the responses of the total sample as a function of question wording. As expected, more respondents endorsed the "not allow" than the "forbid" statement and the "not forbid" than the "allow" statement, $\chi^2(1) = 23.2, p < .001$. Thus, the data replicate the forbid-allow asymmetry resulting in a wording effect of 16.5 percentage points.

Separate analyses of the responses provided by respondents who

Table 1. The Forbid–Allow Asymmetry as a Function of Respondents' Attitude

Forbid Form		*Allow Form*	
Do you think that one should generally forbid the use of salt?		Do you think that one should generally allow the use of salt?	
Yes (forbid)	62.3% (215)	No (not allow)	78.8% (283)
No (not forbid)	37.7% (130)	Yes (allow)	21.2% (76)
	100 (345)		100 (359)
Response × form $\chi^2 = 23.2$, df = 1, $p < .001$, Don't know: $N = 16$			
Attitude Opposed to the Use of Salt			
Yes (forbid)	86.4 (186)	No (not allow)	90.7 (215)
No (not forbid)	13.6 (29)	Yes (allow)	9.3 (22)
	100.0 (215)		100.0 (237)
Response × form $\chi^2 = 2.0$, df = 1, n.s.			
Attitude in Favor of the Use of Salt			
Yes (forbid)	25.8 (17)	No (not allow)	38.7 (24)
No (not forbid)	74.2 (49)	Yes (allow)	61.3 (38)
	100.0 (66)		100.0 (62)
Response × form $\chi^2 = 2.5$, df = 1, n.s..			
Attitude Indifferent to the Use of Salt			
Yes (forbid)	18.8 (12)	No (not allow)	73.3 (44)
No (not forbid)	81.2 (52)	Yes (allow)	26.7 (16)
	100.00 (64)		100.0 (60)
Response × form $\chi^2 = 37.3$, df = 1, $p < .001$			

had reported a pro, con, or indifferent attitude toward the use of castor salt are presented in the lower part of Table 1. A total of 89 percent of the respondents who reported an opposed attitude toward castor salt also wanted the use of castor salt to be discontinued. As expected, their responses were *not* significantly affected by question wording $\chi^2(1) = 2.0$, n.s., and only a small effect of 4.3 percentage points was obtained. Similarly, no significant wording effect was obtained for respondents in favor of the use of castor salt, $\chi^2(1) = 2.5$, n.s., though the difference of 13.1 percentage points was somewhat larger than for respondents holding an opposed attitude. Moreover, 32 percent of the respondents who had reported a favorable attitude on the earlier question nevertheless wanted the use of castor salt to be discontinued. This might be due to the information about its negative effects on plants and trees provided in the question text, which was not provided in the earlier attitude question.

Finally, a large and highly significant wording effect of 54.5 percentage points, $\chi^2(1) = 37.3$, $p < .001$ was obtained for respondents who had reported an indifferent attitude. Specifically, these individuals were likely to respond no to *both* the "allow" (81.2 percent) and the "forbid" (73.3 percent) form of the question, as suggested by our theoretical analysis. Thus, the present data suport the hypothesis that the forbid-allow asymmetry is primarily due to the response behavior of indifferent respondents who will neither endorse that something should be allowed nor that it should be forbidden, because they want to express neither support nor opposition.

While these results are in line with our theoretical analysis, a methodological issue requires additional consideration. To identify indifferents, respondents' attitude toward the use of salt was assessed independently of and prior to the experimental question. Though the questions were separated by 23 unrelated ones, this raises the possibility of consistency effects. Most important, respondents who first reported being indifferent about the issue may be particularly unlikely to endorse a "strong," i.e., affirmative, option later on. Thus, a consistency effect might increase the size of the forbid-allow asymmetry for indifferents. While the present data do not allow us to rule out this possibility, it seems unlikely that a consistency effect alone could account for the pattern of results obtained. Moreover, the fact that 32 percent and 11.3 percent of the respondents who initially reported a favorable or an opposed attitude, respectively, gave an opposite response later on suggests that consistency pressures could not have been very strong.

General Discussion

We have reported here evidence for the stability of the forbid-not allow and allow-not forbid asymmetries across cultures and languages by replicating the effect in Germany (Experiments 1 and 2). Moreover, we attempted to demonstrate that these asymmetries are due to the behavior of respondents who are relatively indifferent about the issue under consideration (Experiment 4). Specifically, we suggested that respondents holding a strong attitude in favor of or in opposition to the issue will give consistent responses independent of question wording. Indifferent respondents, on the other hand, may refuse to endorse *either* that something should be allowed or that it should be forbidden because they are neither strongly in favor of nor strongly opposed to the issue. Therefore, they will be lumped together with respondents holding a strong attitude, thus resulting in the asymmetries under study, with more respondents endorsing the negative than the affirmative form of each question.

Two factors may contribute to the response behavior of indifferents. Most important, respondents are likely to focus on the implications of *doing* what they are asked about (that is, allowing or forbidding something) rather than not doing it, thus exhibiting a well-documented bias to focus on positive instances of a behavior (cf. Fazio, et al., 1982; Nisbett and Ross, 1980). Therefore, respondents may miss the implications of their negative response, feeling that they only gave their opinion about allowing (forbidding) something without ever commenting on the opposite. Moreover, the negative wordings are weaker statements than the affirmative wordings (Experiment 3) and may thus be easier to endorse for indifferents.

The present explanation of the forbid-allow asymmetries implies that the effect should be more pronounced in aggregated data the larger the number of indifferents. Therefore, the effect should be the larger the more abstract the issue, the less salient the issue, and the lower the education of the respondents. Indeed, the data reviewed by Schuman and Presser (1981) provide evidence in support of each of these variables. In addition, the present explanation suggests that the forbid-allow asymmetries may vanish if indifferent respondents are provided the possibility of expressing their indifference. To the best of our knowledge such a study has not been conducted and the "don't know" option provided in some of Schuman and Presser's (1981) studies seems inappropriate to attain this goal. If respondents do indeed focus on the behavior described in the question, that is, to forbid or to allow, indifferents *do know* that they do not want to forbid or to allow something. Thus, responding "don't know" would not accommodate their need to express indifference. "Don't care" or "leave things as they are" categories might be more appropriate to identify indifferents, and to avoid their being lumped together with respondents holding an attitude they do not share.

References

Fazio, R. H., S. J. Sherman, and P. M. Herr
1982 "The feature positive effect in the self-perception process: does not doing matter as much as doing?" Journal of Personality and Social Psychology 42:404–11.
Kalton, G., and Howard Schuman
1980 The effect of the question on survey responses: a review. Proceedings of the Section on Survey Research Methods, Washington, D.C.: American Statistical Association.
Nisbett, R. and Lee Ross
1980 Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall.
Noelle-Neumann, E.
1980 Die Schweigespirale. München: Piper.
Rosenthal, R.
1978 "Combining results of independent studies." Psychological Bulletin 85:185–93.

Rugg, D.
1941 "Experiments in wording questions: II." Public Opinion Quarterly 5:91–92.
Schuman, H., and S. Presser
1981 Questions and Answers in Attitude Surveys. New York: Academic Press.

Appendix A

German text of the questions used in the reported experiments

EXPERIMENT 1

Meinen Sie, der Gesetzgeber sollte die sogenannten Peep-Shows, also das bezahlte Sichzurschaustellen nackter Frauen in aufreizenden Posen, generell verbieten (erlauben)? (Ja/Nein)

EXPERIMENT 2

Meinen Sie, der Gesetzgeber sollte die Vorführung von pornographischen Filmen in öffentlichen Kinos generell verbieten (erlauben)? (Ja/Nein)

EXPERIMENT 3

Wir hätten gerne von Ihnen eine Beurteilung der folgenden Aussage. Herr Müller meint: "Der Gesetzgeber sollte die sogenannten Peep-Shows, also das bezahlte Sichzurschaustellen nackter Frauen in aufreizenden Posen, generell verbieten (erlauben/nicht erlauben/nicht verbieten)."

Wie sehr ist Ihrer Ansicht nach Herr Müller für bzw. gegen Peep-Shows? Machen Sie bitte ein Kreuz auf der folgenden Skala.

+5	+4	+3	+2	+1	0	−1	−2	−3	−4	−5

extrem *für*
Peep-Shows

extrem *gegen*
Peep-Shows

EXPERIMENT 4

Q.10. Wir hätten gerne von Ihnen gewußt, wie Sie zu den folgenden Aussagen stehen. Stimmen Sie der jeweiligen Aussage zu, ist sie Ihnen egal oder lehnen Sie sie ab?
Streusalz sollte bei Glatteis nicht mehr verwendet werden (ja, stimme zu; ist mir egal/weiß nicht; nein, lehne ab)

Q.34. Durch das Streuen von Salz im Winter sind Pflanzen und Bäume in Straßennähe gefährdet. Sollte man das Salzstreuen generell erlauben (verbieten)? (Ja/Nein)

Interviewers' Voices and Refusal Rates in Telephone Surveys

LOIS OKSENBERG, LERITA COLEMAN
AND CHARLES F. CANNELL

As SURVEY RESEARCH uses telephone interviews more extensively, new methodological problems become apparent. One problem is response rates, which are generally lower than for face-to-face interviews. The reasons for the lower response rates are not known. Refusals in face-to-face interviews are often attributed not only to the content but to the respondent's perception of the interviewer during initial introductory phrases and impressions based on his or her appearance. In phone interviews these latter cues are absent, and auditory stimuli are the sole basis for evaluation. This implies that characteristics of voice and speech patterns provide major bases for first impressions, and may relate to acceptance or rejection of the interview.

In listening to telephone interviewers at work one is impressed with their unique speech characteristics. Further, interviewers are characterized by high or low response rates, and they tend to maintain their

---

Abstract Interviewers' voice characteristics are related to refusal rates in telephone interviews. Judges reliably rated vocal and personal characteristics of telephone interviewers from recordings of the first one-half minute of interview introductions. Interviewers with low refusal rates were judged as having comparatively higher pitched voices, greater ranges of variation in pitch, greater loudness, faster rates of speaking, and clearer and more distinct pronunciation. They also were judged as more competent and as having a more positive approach to the respondent and the interview.

The authors are with the Institute for Social Research of The University of Michigan: Lois Oksenberg is a Research Investigator and Charles F. Cannell is Program Director Emeritus in the Survey Research Center; Lerita Coleman is an Assistant Research Scientist in the Research Center for Group Dynamics. The authors are grateful to Maryellen McSweeney for her statistical advice. This material is based upon work supported by the National Science Foundation under Grant No. SES-8112801.

Public Opinion Quarterly Vol. 50:97–111 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-97/$2.50

rate over time. For example, Survey Research Center (The University of Michigan) telephone interviewers working on a continuing monthly survey showed considerable consistency in refusal rates over time, with rates ranging from 6 to 42 percent.

These observations led us to attempt a systematic analysis of interviewers' voice characteristics to see whether we could identify factors correlated with refusal rates. If factors could be identified, we might be able to select interviewers with positive voices or train interviewers to speak more acceptably.

Of the nonresponses in telephone surveys, one-half to two-thirds are refusals to be interviewed. The great majority of refusals occur within the opening minutes of the interview, prior to the first substantive question. In a study at the Survey Research Center (SRC), for example, about a third of the refusals occurred in the opening seconds, another third occurred during the body of the introduction, while the remainder occurred at the point of listing household members prior to the first question. Respondents appear to form an attitude toward the interview early and accept or refuse to participate accordingly.

Efforts to improve response rates have focused on the content of introductions, but have met with little success (Dillman et al., 1976; O'Neil et al., 1979). Differences in refusal rates among interviewers appear to reflect factors unrelated to content. Listening to interview introductions reinforced the idea that interviewers' vocal characteristics are key to the impressions potential respondents form of them and whether they would consent to and would enjoy an extended interview with them.

Initially a person probably will have a somewhat negative attitude toward a telephone call from a stranger. The caller is likely to be selling something, soliciting contributions, making a political appeal, or imposing on the recipient in some other way. As conversation progresses, the recipient may change his or her attitude toward the call. If the purpose of the call is attractive, he or she may become less negatively or even positively disposed. If the purpose is unattractive or unclear, the recipient will become more negative. The characteristics exhibited by the caller will also affect the recipient's attitude. There will be less interest in engaging in extended conversation with an unappealing caller.

The telephone interviewer shares these barriers to communication. The presentation must be sufficiently positive to overcome initial indifference or negative feelings. In a face-to-face interview there are a number of cues to help establish a positive attitude: The interviewer's vocal characteristics, physical appearance, dress, demeanor, interviewing paraphernalia and, perhaps, motor vehicle are all evident. In contrast, a telephone interviewer must create a sufficiently positive impression with only vocal cues. If vocal characteristics lead the respondent to perceive the interviewer as unappealing, cooperation will be less likely.

It is evident from the history of research on paralanguage and language attitudes that listeners rely heavily on vocal characteristics to form impressions and attitudes about speakers (Kramer, 1963; Giles and Powesland, 1975; Giles et al., 1980). Judges tend to agree in rating speakers' personal attributes from their voices (Addington, 1968; Allport and Cantril, 1934; Kramer, 1964). While rater agreement appears to reflect vocal stereotypes rather than accurate judgments, it nonetheless indicates that people hold definite ideas about the correlation of vocal attributes with personality.

Judgments about demographic characteristics are based on vocal cues. Sex, age, social status, and race are accurately identified from the voice (Smith, 1979; Pear, 1931; Harms, 1961). These demographic characteristics appear to influence judgments of personal characteristics. Low ascribed status, for example, is associated with less credibility (Harms, 1961; Putnam and O'Hearn, 1955; Ellis, 1967). Recent research has noted that, in general, female speakers are perceived to be more socially attractive and competent than male speakers (Smith, 1979).

A number of studies, primarily using male speakers, have attempted to isolate and identify specific vocal cues involved in evaluation. These studies have utilized a number of vocal features including pitch, rate of speech, amplitude (loudness), intonation (variation in pitch), pauses, and fluency. Increases in pitch, for example, produce judgments of less competence and less benevolence, while faster speech elicits ratings of competence (Brown et al., 1973). Greater variability in pitch, more speech volume, increased speech rate and less halting speech have been linked to persuasiveness (Mehrabian and Williams, 1969; Packwood, 1974). Rapid speech rates elicit ratings of more credibility and enhanced persuasiveness (Pearce and Conklin, 1971). Faster speakers are rated as more intelligent, knowledgeable, and objective than slower speakers (Miller et al., 1976). Perceived persuasiveness may not result from a fast rate of speech per se but from a smooth, nonhesitant speaking style. Large numbers of pauses in speech produce negative ratings of speakers (e.g., as angry, anxious, contemptuous, and bored [Lalljee, 1971]). Other highly desirable traits, including credibility, are associated with nonhesitant speakers (Lay and Burron, 1968).

The effect of vocal characteristics on personal attributions may depend on the sex of the speaker. Addington (1968) found that greater variation in pitch was interpreted somewhat differently for male and female speakers. To further complicate matters, Apple et al. (1979) found amplitude, pitch, and rate of speaking to be intercorrelated in normal speech, making it difficult to link particular characteristics to perceptions of the speaker.

The literature on vocal cues prompted a number of specific questions about the effects of voices of telephone interviewers. What cues do respondents use to judge interviewers? Is there agreement on the characteristics of voices and on the personal attributes signalled by them, or are these perceptions so idiosyncratic that no general patterns emerge? If there is agreement on the voice and personal attributes signalled by voices, are these perceptions related to interviewer refusal rates?

Three objectives for this study were derived from these questions: (1) Can vocal qualities of telephone interviewers be judged reliably? (2) Do ratings of vocal qualities correlate with perceptions of desirable and undesirable personal attributes? (3) Is there a relation between these characteristics and rate of refusals?

This article reports our first investigations of these issues. In two studies, raters listened to recordings of the first one-half minute of telephone interview introductions and evaluated a number of the interviewers' vocal and personal characteristics. Because of the small number of male interviewers on the SRC staff and the need to restrict the scope of the research, only female interviewers were included. In addition to the rating studies, acoustical analyses of physical properties of interviewer voices were performed and reported by Sharf and Lehman (1984).

Measures of Vocal and Personal Characteristics

Rating scales were developed by the project staff on the basis of previous paralinguistic, personality, person-perception, and attitude change studies (Figures 1 and 2). Our goal was to represent key vocal cues and personal attributes identified in the literature, as well as characteristics that, on the basis of our own evaluations, seemed particularly relevant to telephone interviewing. We included only the number of vocal or personal characteristics that could be rated in a single session. After extensive investigation, 10 scales measuring vocal characteristics and 21 scales measuring personal attributes were selected. With the exception of a five-category scale for the interviewer's social class and a six-category scale for age, all scales were seven-point anchored semantic differential type.

Figure 1. Rating Interviewer's Voice Quality

VOICE NUMBER ○

Please rate the voice on the following scales.
For each scale, circle the number that best describes what you hear.

Average pitch	low	1 2 3 4 5 6 7	high
Variation in pitch (intonation)	small range	1 2 3 4 5 6 7	large range
Loudness	soft	1 2 3 4 5 6 7	loud
Stress or emphasis of words and phrases	done appropriately	1 2 3 4 5 6 7	not done appropriately
Flow of words	smooth and easy	1 2 3 4 5 6 7	hesitant
Pronunciation of words	clear and distinct	1 2 3 4 5 6 7	indistinct or slurred
Spontaneity	sounds natural/ spontaneous	1 2 3 4 5 6 7	sounds memorized or like reading or otherwise "canned"
Rate of speaking	slow	1 2 3 4 5 6 7	fast
Any sort of accent? Please check: ____ NO ____ YES. IF YES:	pleasant	1 2 3 4 5 6 7	unpleasant
Generally, how pleasant was this voice to listen to:	pleasant	1 2 3 4 5 6 7	unpleasant

Recordings

Introductory statements made by experienced SRC interviewers were recorded. The recordings of actual telephone interviews were made through an extension phone in order to reproduce as closely as possible the voice as heard by the respondent. Interviewers used the standardized introduction provided for the survey with little variation. The recorded portion read:

> Hello, my name is ______. I'm calling from The University of Michigan, in Ann Arbor. Here at the university we are currently working on a study for the Survey Research Center. First of all I need to be sure I've dialed the right number. Is this ______ (telephone number)? Since this telephone number has been generated by a computer, I do not know whether this number is for a business or a home. Which is it?

The recordings were approximately 30 seconds long. Interviewers' names and respondents' area codes were deleted from recorded introductions.

Figure 2. Rating Interviewer's Personal Characteristics

Interviewer Number ○

Please rate the interviewer's manner and personal characteristics on the following scales. For each scale, circle the number that best describes your perception of the interviewer.

young	1 2 3 4 5 6 7	old
cheerful	1 2 3 4 5 6 7	uncheerful
good at her job	1 2 3 4 5 6 7	poor at her job
interesting	1 2 3 4 5 6 7	boring
self-assured	1 2 3 4 5 6 7	ill at ease
educated	1 2 3 4 5 6 7	uneducated
trustworthy	1 2 3 4 5 6 7	untrustworthy
friendly	1 2 3 4 5 6 7	unfriendly
intelligent	1 2 3 4 5 6 7	unintelligent
enthusiastic	1 2 3 4 5 6 7	apathetic
impressive	1 2 3 4 5 6 7	unimpressive
interested in her task	1 2 3 4 5 6 7	bored with her task
polite	1 2 3 4 5 6 7	impolite
professional	1 2 3 4 5 6 7	amateurish
kind	1 2 3 4 5 6 7	unkind
mature	1 2 3 4 5 6 7	immature
pleasant	1 2 3 4 5 6 7	unpleasant
good-looking	1 2 3 4 5 6 7	plain

What do you think the interviewer's social class is? (Please check)

____ working
____ lower-middle
____ middle
____ upper-middle
____ upper

What do you think the interviewer's age is? (Please check)

____ under 20
____ 20–29
____ 30–39
____ 40–49
____ 50–59
____ 60–69

How would you like to be interviewed *by this person?*

like to be interviewed by her	1 2 3 4 5 6 7	dislike to be interviewed by her

The Two Studies

Study 1 was based on recordings of six interviewers; three interviewers with low refusal rates (.03, .11, .14) and three with high refusal rates (.36, .38, .46) over a seven-month period were selected. A seventh interviewer with an intermediate refusal rate was included for practice ratings.

Eighteen students (11 women and 7 men) in a summer course in survey methodology at The University of Michigan served as raters. They ranged in age from the twenties to the fifties. Ratings of the vocal characteristics for each of the six voices were obtained in one session, and ratings of the personal characteristics for the same six voices were obtained in another session the following day. This minimized influence between ratings of the two types of characteristics and avoided too great a burden on raters.

On the first day, after a brief introduction to their task, raters heard the practice voice twice, and then had about one minute for rating. The same procedure was followed for the remaining voices. Voices of the three more successful interviewers and the three less successful interviewers were alternated. The procedure in the second session, in which personal characteristics were rated, followed the same pattern.

A second study (Study 2) used raters recruited from a local church. The raters ranged in age from 18 to over 60 and were primarily working class or middle class and predominantly white men and women (22 females and 18 males). As in the first study, ratings were obtained for six interviewers: three with low refusal rates (.10, .13, and .14) and three with high refusal rates (.24, .29, and .38). Two of the six voices and the practice voice had been included in the first study. Procedures were the same as those in Study 1, except that the two rating sessions were on consecutive Sundays, and the raters received an introduction to the task better geared to a general audience.

Construction of Measures and Reliability

Rating scale reliabilities were obtained for each of the studies. The analysis of variance method with repeated measures giving mean interrater reliabilities as described by Winer (1971) was applied. Reliabilities of the mean ratings assigned by the listeners were generally very high—the great majority in the .80s and .90s in both studies. High correlations among many of the ratings, however, indicated considerable overlap among scales. Ratings for four vocal characteristics were so highly intercorrelated in both studies that they appeared to measure a single factor. They included *stress or emphasis of words and phrases, flow of words, spontaneity,* and *generally, how pleasant was this voice to listen to.* This conclusion was supported by principal axes factor analyses of the ratings of vocal characteristics undertaken separately for the two studies. Since the object was to investigate relations among ratings made by individual listeners, each listener in a study contributed six sets of ratings to the analysis, one set for each voice. Ratings of the four characteristics were combined in one measure referred to as *speaking skill,* scored as the mean of the four ratings. The remaining vocal characteristics showed lower levels of correlation and were kept separate. The scale measuring *pleasantness of accent* was not analyzed because of the small number of times interviewers were rated as having accents.

Personal characteristic ratings showed generally high levels of correlation in both studies as well as some grouping. For all personal characteristics (except *young, social class,* and *age category*) one anchor of the scale was positively valued, the other negatively valued. To explore how best to summarize findings concerning these evaluative scales, we performed principal axes factor analyses including only the 15 evaluative scales that were reliable in both studies.[1] The *polite, kind,* and *trustworthy* scales were not included in these or subsequent analyses. They were quite unreliable in Study 1 since all six voices were rated uniformly high. For each study one factor including all 15 scales provided an economical description of their interrelations. (A single factor accounted for 66 percent of the variance in Study 1, 57 percent in Study 2.) Accordingly, a *general attractiveness* measure was constructed and scored as the mean of ratings on the 15 scales.

Correlations among the 15 scales, however, showed some clustering into two groups: one related to competence of the interviewer, the other to her approach to the respondent and the interview. The theoretical interest of these groups of scales led us to perform additional factor analyses to identify relatively pure measures of the underlying concepts. For each study we extracted two factors and rotated them using an oblique solution. Scales that had very high loadings on the first factor and much lower loadings on the second factor were grouped together, while scales with very high loadings on the second factor and much lower loadings on the first factor likewise were grouped. Groups formed for each of the studies were then compared. Scales grouped together in both studies were used to measure the relevant concepts. The first concept measured was *competence* including scales for *educated, intelligent,* and *professional*. The second concept measured was *positive approach* including scales for *cheerful, friendly, enthusiastic, interested in her task,* and *pleasant*. Both measures were scored as the mean of the ratings on the respective scales.

Table 1 gives the mean inter-rater reliabilities for the measures. The *young* scale was not analyzed because it yielded results indistinguishable from *age category*. The extremely high levels of reliability show that listeners reliably rated vocal and personal characteristics of telephone interviewers from recordings of the first 30 seconds or so of their introductions.[2]

---

[1] To control for differences in listeners' use of the scales, the six ratings for each characteristic contributed by a given listener were standardized by subtracting each rating from the mean of the six.

[2] Levels of agreement between any two raters are lower than the reliabilities reported here. For example, a mean inter-rater reliability of .91 in Study 1 corresponds to an estimated mean correlation between ratings of all pairs of listeners of .35 in Study 1 and of .20 in Study 2 (Rosenthal, 1973). Individuals particularly skilled in making the judgments must be identified or trained in order to obtain reliable ratings from only one or a very few raters.

Table 1. Mean Inter-Rater Reliabilities of Interviewer Characteristic Measures in Two Studies

Scale	*Study 1*[a]	*Study 2*[b]
Vocal characteristics		
Average pitch	.95	.98
Variation in pitch	.88	.89
Rate of speaking	.90	.88
Loudness	.86	.93
Pronunciation	.88	.86
Speaking skill[c]	.92	.89
Personal characteristics		
General attractiveness[d]	.91	.89
Positive approach[e]	.90	.90
Competence[f]	.89	.86
Social class	.83	.78
Age category	.95	.99

[a] Figures are based on 18 raters (except that for *pronunciation*, which is based on 17 raters).

[b] Figures are based on 40 raters (except that for *loudness*, which is based on 36 raters, and those for *variation in pitch* and *rate of speaking*, which are based on 39 raters).

[c] Includes *stress and emphasis on words and phrases*, *flow of words*, *spontaneity*, and *generally, how pleasant was this voice to listen to*.

[d] Includes 15 personal characteristics: those in the *positive approach* and *competence* measures, plus *good at her job*, *interesting*, *self-assured*, *impressive*, *mature*, *good looking*, and *like to be interviewed by her*.

[e] Includes *cheerful*, *friendly*, *enthusiastic*, *interested in her task*, and *pleasant*.

[f] Includes *educated*, *intelligent*, and *professional*.

Relation of Rated Characteristics to Refusal Rates

Tables 2 and 3 give the mean ratings for the three interviewers with low refusal rates and for the three interviewers with high refusal rates in each of the two studies. The tables also give results of comparisons of the mean ratings for the two groups of interviewers, one comparison for each measure. The comparisons were tested within the context of the respective analyses of variance used to determine reliabilities. The tables also give η^2, which is the proportion of variation in the ratings attributable to the difference between the two groups of interviewers. Since the significance levels for η^2 are the same as those of the F for the associated comparison, they are noted only for the F values.

The η^2 for all vocal measures differed significantly from zero in both studies, and several were quite large (Table 2). In Study 1 the two elements of pitch (*average pitch* and *variation in pitch*) were strongly related to interviewer refusal rates, with η^2 of .30 and .19, respectively. They were followed by *loudness*, *pronunciation*, *rate of speaking*, and *speaking skill*. The η^2 in Study 2 were somewhat lower. *Loudness* had the highest value (.15), followed by *average pitch*, *variation in pitch*,

Table 2. Mean Ratings of Interviewers Grouped by Refusal Rates, *F*s, η^2, and Direction of Differences for Vocal Characteristics in Two Studies

	Mean Rating: Interviewer Group				
Characteristic	*Low Refusal Rate*	*High Refusal Rate*	*F**	η^2	*Characterization of Interviewers with Low Refusal Rates*
Average pitch	4.76	3.24	64.19^a	.30	Higher
	4.63	*3.66*	*58.46*^{b}	*.11*	pitch
Variation	4.61	3.15	29.83	.19	Greater
in pitch	*4.62*	*3.90*	*24.74*	*.08*	variation
Loudness	4.65	3.76	28.56	.17	Louder
	4.80	*4.06*	*59.90*	*.15*	
Rate of speaking	4.20	3.56	14.93	.07	Faster
	4.01	*3.39*	*24.35*	*.08*	
Pronunciation	5.19	4.23	19.19	.11	Clearer and
	5.69	*4.98*	*21.82*	*.06*	more distinct
Speaking skill	4.22	3.67	7.91	.04	More
	4.62	*4.25*	*6.60*^{c}	*.02*	skilled

NOTE: Measures were scored as indicated in Figure 1, except for *pronunciation* and *speaking skill*. Scoring was reversed for these two measures so that high scores indicated clear and distinct pronunciation and good speaking skill. Italicized figures are for Study 2. Other figures are for Study 1.

* Except as noted $p < .01$

[a] For five characteristics in Study 1 $df = 1,85$. For one characteristic $df = 1,80$.

[b] For three characteristics in Study 2 $df = 1,195$. For three others $df = 1,190$ or $1,185$.

[c] $p < .05$

rate of speaking, *pronunciation*, and *speaking skill*. In both studies, interviewers with lower refusal rates were rated as having higher pitched voices, greater ranges of variation in pitch, greater loudness, faster rates of speaking, and clearer and more distinct pronunciation. *Speaking skill* was only weakly related to refusal rates.

The η^2 for all personal characteristics also differed significantly from zero (Table 3). The largest in Study 1 was for *positive approach* (.21), followed by *general attractiveness*, *age category*, *social class*, and *competence*. The η^2 in Study 2 again were somewhat lower. *General attractiveness*, *competence*, and *age category* had equally high values (.10), followed by *positive approach* and *social class*. In both studies interviewers judged as more generally attractive and, more specifically, as having a more positive approach and greater competence, had lower refusal rates. Interviewers perceived to have higher social class also had lower refusal rates.

Although age was related to interviewer success in both studies, the results were contradictory. It happened that each study included one interviewer noticeably older than the others. In Study 1 she was one of

Table 3. Mean Ratings of Interviewers Grouped by Refusal Rates, Fs, η^2, and Direction of Differences for Personal Characteristic Measures in Two Studies

	Mean Rating: Interviewer Group				
Characteristic	*Low Refusal Rate*	*High Refusal Rate*	F^*	η^2	*Characterization of Interviewers with Low Refusal Rates*
General	4.70	3.86	28.06[a]	.15	More
attractiveness	*4.89*	*4.28*	*38.15*[b]	*.10*	attractive
Positive	5.00	3.93	34.96	.21	More
approach	*4.86*	*4.29*	*26.52*	*.07*	positive
Competence	4.67	4.01	16.12	.08	More
	5.17	*4.48*	*41.22*	*.10*	competent
Social	3.06	2.56	17.82	.10	Higher social
class	*3.29*	*2.98*	*17.02*	*.05*	class
Age	3.11	2.52	23.84	.11	Younger
category	*2.65*	*3.38*	*110.45*	*.10*	Older

NOTE: Scoring for the *general attractiveness*, *positive approach*, and *competence* measures was reversed from that indicated in Figure 2 for the scales included in them. High scores on these measures indicate positive evaluation. *Social class* was scored so that WORKING = 1 and UPPER = 5. *Age category* was scored so that UNDER 20 = 1 and 60–69 = 6. Italicized figures are for Study 2; other figures are for Study 1.

* For all measures $p < .01$

[a] For characteristics in Study 1 df = 1,85.

[b] For characteristics in Study 2 df = 1,195.

the more successful interviewers, while in Study 2 she was one of the less successful interviewers. We think the results for age reflect vocal and personal characteristics associated with interviewing success, rather than factors related to age per se.

Relation of Vocal Characteristics to Personal Attributions

There was generally a high level of correlation in both studies between mean ratings of vocal and of personal characteristics for the six interviewers (Table 4). Of the vocal characteristics, *variation in pitch* had the highest correlations with interviewer attractiveness, whether measured by *general attractiveness* or two of its components, *positive approach* or *competence*. The other vocal measures also had substantial correlations with the attractiveness measures. Thus, greater attractiveness was associated with voices perceived as relatively high-pitched and loud, and as having a large range of variation in pitch, a fast rate of speaking, clear and distinct pronunciation, and good speaking skills.

Perceptions of higher social class were associated with voices having

Table 4. Correlations Between Mean Ratings of Vocal and Personal Characteristic Measures in Two Studies

	Personal Characteristics				
Vocal Characteristics	*General Attractiveness*	*Positive Approach*	*Competence*	*Social Class*	*Age Category*
Average pitch	.61	.78	.48	.58	.75
	.56	*.56*	*.61*	*.63*	*−.93*
Variation in pitch	.95	.98	.87	.85	.40
	.77	*.60*	*.94*	*.96*	*−.75*
Loudness	.44	.49	.40	.63	.18
	.67	*.47*	*.91*	*.92*	*−.61*
Rate of speaking	.76	.73	.70	.48	−.15
	.89	*.84*	*.71*	*.49*	*−.15*
Pronunciation	.88	.84	.89	.94	.59
	.48	*.41*	*.64*	*.65*	*−.65*
Speaking skill	.86	.68	.93	.84	−.04
	.53	*.45*	*.36*	*.19*	*.33*

NOTE. Figures in italics are for Study 2; other figures are for Study 1. All are product moment coefficients based on six interviewers. Measures were scored as described in the notes to Tables 2 and 3.

the same characteristics. Results for age again were contradictory in the two studies, most likely for the reason already noted.

Discussion

All three research goals were met: reliable ratings of vocal and personal characteristics were obtained; vocal characteristics were correlated with desirable and undesirable personal characteristics; and both were related to interviewer refusal rates. In both studies higher pitch, greater variation in pitch, loudness, faster rate of speaking, clearer and more distinct pronunciation, and good speaking skills were all associated with attractiveness (whether determined by the *general attractiveness* measure or by *positive approach* or *competence*) and with higher perceived social class. Interviewers with these vocal and personal characteristics had low refusal rates.[3]

While the associations of vocal and personal characteristics with refusal rates were highly significant statistically, in most instances they were only moderately strong. The interviewers in these studies were a select group, since screening had already eliminated job applicants with voices judged unsuitable, and early departure of some apparently less skilled interviewers excluded them from the studies. A more heterogeneous set of interviewers might yield stronger findings.

---

[3] While a fast rate of speaking for introductory comments may help response rates, in asking questions a slower pace is probably better because it facilitates comprehension and attention to reporting tasks.

Except for the desirability of high average pitch, the findings are consistent with expectations from paralinguistic studies. In these studies (primarily using male speakers) higher pitch was rarely associated with positive attributes. Given gender expectations, higher pitched voices are more acceptable for females just as lower pitched voices are expected of males. It may well be that for females, high pitch (within an acceptable range) is more attractive than low pitch. Although acceptable ranges have not been established, especially high-pitched female voices may not be evaluated positively.

The finding that ratings of loudness were related to success is puzzling. Whereas other acoustical properties of the voices were reproduced accurately in the recordings, vocal intensity was not. The intensities as recorded were roughly the same for all interviewers. The procedure, however, preserved the relative intensities of the various portions of each recording. We can only speculate whether such intensity patterns or some other acoustic properties of the recordings led raters to judge some voices as louder than others.

Findings concerning personal attributions also raise questions. Is there benefit in separating use of a positive approach from competence, or should they be subsumed in a general measure of interviewer attractiveness? Although *positive approach* and *competence* were about equally strongly related to refusal rates in Study 2, *positive approach* was considerably more strongly related than *competence* to refusal rates in Study 1. The concepts are of particular interest since they resemble the task skills and social-emotional dimensions identified in studies of leadership in small groups (Bales, 1966). Further study of them is warranted.

Our studies were limited to female interviewers who were fairly homogeneous with regard to perceived education, social class, age, accent, and use of standard pronunciation. Research is needed to explore the effects of departures from these characteristics as well as to explore whether desirable characteristics differ for male and female interviewers. With regard to the vocal characteristics investigated here, optimal levels have yet to be established. Not only is very high pitch likely to be undesirable, but also a very large range of variation in pitch, great loudness, and an extremely fast rate of speaking.

Finally, we have assumed that vocal characteristics lead to an assessment of the interviewer that in turn disposes the respondent to cooperate or not. However, in our studies vocal characteristics were more strongly related to refusal rates than were personal characteristics. It may be that vocal characteristics have a direct impact on the listener that is not mediated by impressions of the speaker as a person. More research is needed to clarify the process by which these characteristics influence cooperation.

This investigation reinforces the hypothesis that respondents react to cues communicated by the interviewer's voice and may grant or refuse an interview on this basis. Content of the introduction appears to be less important than impressions and affects aroused by the vocalization. To paraphrase a line from *My Fair Lady*, "Respondents don't care what you say, actually; it's whether you say it properly." This should not be surprising to survey researchers, who have always stressed the importance of personal appearance and approach to respondents on the doorstep. Vocal cues, as the only source of information about telephone interviewers, play a particularly significant role in telephone interviewing.

This study should be viewed principally as emphasizing the importance of paralinguistic research to survey methodologists. We hope that other researchers will investigate paralinguistic issues, attempt to replicate our findings, and explore their implications in greater detail. The findings also introduce the question as to whether voice characteristics should be a selection criterion or whether voice training should be added to interviewer training.

References

Addington, D. W.
1968 "The relationship of selected vocal characteristics to personality perception." Speech Monographs 35:492–503.
Allport, G., and H. Cantril
1934 "Judging personality from voice." Journal of Social Psychology 5:37–54.
Apple, W., L. A. Streeter, and R. M. Krauss
1979 "Effects of pitch and speech rate on personal attributions." Journal of Personality and Social Psychology 37:715–26.
Bales, R. F.
1966 "The equilibrium problem in small groups," in A. P. Hare, E. F. Borgatta, and R. F. Bales (eds.), Small Groups: Studies in Social Interactions. New York: Alfred A. Knopf.
Brown, B. L., W. L. Strong, and A. C. Rencher
1973 "Perceptions of personality from speech: effects of manipulations of acoustical parameters." Journal of Acoustical Society of America 54:29–35.
Dillman, D., J. G. Gallegos, and J. H. Frey
1976 "Reducing refusal rates for telephone interviews." Public Opinion Quarterly 40:66–78.
Ellis, D.
1967 "Speech and social status in America." Social Forces 45:431–37.
Giles, H., and P. F. Powesland
1975 Speech Style and Social Evaluation. New York: Academic Press.
Giles, H., W. P. Robinson, and P. M. Smith (eds.)
1980 Language: Social Psychological Perspectives. New York: Pergamon.

Harms, L. S.
1961 "Listener's judgments of status cues in speech." Quarterly Journal of Speech 47:164–68.
Kramer, E.
1963 "Judgment of personal characteristics and emotions from nonverbal properties of speech." Psychological Bulletin 60:408–20.
1964 "Personality stereotypes in voice: a reconsideration of the data." Journal of Abnormal and Social Psychology 68:247–51.
Lalljee, M. G.
1971 "Disfluencies in normal English speech." Unpublished Ph.D. dissertation, Oxford University.
Lay, C. H., and B. F. Burron
1968 "Perception of the personality of the hesitant speaker." Perceptual and Motor Skills 23:951–56.
Mehrabian, A., and M. Williams
1969 "Nonverbal concomitants of perceived and intended persuasiveness." Journal of Personality and Social Psychology 13:37–58.
Miller, N., G. Maruyana, R. Beaber, and K. Valone
1976 "Speed of speech and persuasion." Journal of Personality and Social Psychology 34:615–24.
O'Neil, M. J., R. M. Groves, and C. F. Cannell
1979 "Telephone interview introductions and refusal rates: experiments in increasing respondent cooperation." American Statistical Association meetings, August.
Packwood, W. T.
1974 "Loudness as a variable in persuasion." Journal of Counseling Psychology 21:1–2.
Pear, T. H.
1931 Voice and Personality. London: Chapman and Hall.
Pearce, W. B., and F. Conklin
1971 "Nonverbal vocalic communication and perceptions of a speaker." Speech Monographs 38:235–41.
Putnam, G. N., and E. M. O'Hearn
1955 "The status significance of an isolated urban dialect." Language 31:1–32.
Rosenthal, R.
1973 "Estimating effective reliabilities in studies that employ judges' ratings." Journal of Clinical Psychology 29:342–45.
Sharf, D. J., and M. E. Lehman.
1984 "Relationship between the speech characteristics and effectiveness of telephone interviewers." Journal of Phonetics 12:219–28.
Smith, P.
1979 "Sex markers in speech," in K. R. Scherer & H. Giles (eds.), Social Markers in Speech. England: Cambridge University Press.
Winer, B. J.
1971 Statistical Principles in Experimental Design. 2nd ed. NY: McGraw-Hill.

The Polls:
ABM and Star Wars:
Attitudes Toward Nuclear Defense, 1945–1985

THOMAS W. GRAHAM AND BERNARD M. KRAMER

PRESIDENT REAGAN'S Strategic Defense Initiative (SDI), popularly called Star Wars, is the most recent manifestation of a 40-year debate over developing systems to defend against nuclear attack. This compilation of national survey research is a selection of questions about nuclear defenses taken from our collection of 61 national and 3 state opinion polls conducted from December 1945 through December 1985.[1]

Despite changes in the US-Soviet nuclear balance, several attitudes toward active defenses against nuclear weapons have been remarkably stable. For example, eight out of ten questions which asked about the generic feasibility of developing a defense against nuclear weapons found a majority (up to two-thirds) of the population believing that the U.S. or its scientists could develop such a defense.[2]

Awareness of both the ABM system and the SDI research program has hovered at the 70 percent level both in the 1969 ABM debate and since Reagan's March 1983 speech.[3] However, no more than 16 percent to 24 percent of the public is following the current Star Wars debate "a lot" or "very closely" (1/85, 2/85). In another area concerning awareness, in 1985 only about one-third of the public has heard about the ABM/SALT I treaty (2/85, 9/85).

---

[1] Neither state polls nor "passive" defenses such as civil defense are covered in this article. For a presentation and analysis of all survey questions and marginals from the 64 surveys, see Thomas W. Graham, Public Attitudes Toward Active Defense: ABM & Star Wars, 1945–1985 (Cambridge, MA: Center for International Studies, MIT, 1986).

[2] These nine polls were conducted in 2/45, 4/46, 8/46, 2/47, 11/49, 6/64, 6/68, 1/85, 5/85. One of the "exceptional" questions asked whether people agreed with expert opinion that it was impossible to defend against surprise attack by rockets carrying atomic bombs (2/45, NORC). The other queried whether the U.S. can "protect" itself from incoming nuclear missiles (5/85, Sindlinger).

[3] Surveys conducted on 4/69, 5/69, 7/69, 4/83, 1/85, 7/85, 9/85.

Thomas W. Graham is a Ph.D. candidate in the Department of Political Science and a Graduate Research Associate at the Defense and Arms Control Studies Program at MIT. Bernard M. Kramer is Professor of Psychology at the University of Massachusetts, Boston.

Public Opinion Quarterly Vol. 50:125–134 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-125/$2.50

Approximately 75 percent of the public holds the (inaccurate) belief that the United States has a fairly effective defense against nuclear weapons (6/64, 6/68, 4/79). In addition, most people are satisfied with our (nonexistent) defense against nuclear attack (5/82, 7/82, 4/84).

In the 1960s, a majority of the population believed that Russia had an existing ABM defense (6/64, 6/68, 4/69). In the 1980s, a plurality of 34 percent thought the Soviets were ahead of the U.S. in Star Wars technology (5/85).

In 1969 deployment of the ABM was favored by the minority of the public who both had heard of the system and had an opinion about building the Safeguard system. In 1971 and 1972, when presented with the idea of spending additional money for the system, support turned to opposition.

A substantial majority of the public has consistently supported the idea of negotiating an ABM agreement or negotiating limitations on Star Wars with the Soviets.[4] Despite this pro-arms control stance, the public believes that moving forward with Star Wars will encourage the Soviet Union to negotiate a nuclear arms control agreement with the U.S. (1/85, 2/85, 5/85).

In the tables that follow, all numbers are percentages. The authors wish to thank the many individuals who helped them discover this wealth of data[5] and the survey organizations that graciously have made it available for academic research:

ABC/WP—ABC & Washington Post
Abt—Clark Abt Associates
AIPO—American Institute of Public Opinion/Gallup
CBS/NYT—CBS & New York Times
Harris—Louis Harris & Associates
LAT—Los Angeles Times
MIS/N—Market Information Service & Dr. Jiri Nehnevaja
M&K—Marttila & Kiley
NORC/N—National Opinion Research Center & Dr. Jiri Nehnevaja
ORC/N—Opinion Research Company & Dr. Jiri Nehnevaja
P&S—Penn & Schoen Associates
Sind—Sindlinger & Associates

ABM

NORC/N & MIS/N [respondents ranked the desirability of building an ABM from +3 to −3]:

	−3	*−2*	*−1*	*0*	*+1*	*+2*	*−3*	*DK*
12/63	8	3	2	5	8	16	59	—
6/64	5	2	1	4	6	13	67	1
2/66	7	1	2	4	7	9	69	1
3/72	6	4	4	14	14	16	43	5

---

[4] Surveys were conducted on 6/68, 8/70, 6/71, 2/72, 3/72, 6/72, 6/73, 11/73, 4/84, 6/84, 7/84, 11/84, 1/85.

[5] Special thanks to the always cooperative Marilyn Potter, Lois Timms and John Benson of the Roper Center for Public Opinion Research at the University of Connecticut; Ann Gray and Josephine Marsh of the Louis Harris Data Center, Institute for Research in Social Sciences at the University of North Carolina; Professor Jiri Nehnevaja and Nicholas R. Trio of the University Center for Social and Urban Research at the University of Pittsburgh; Tom W. Smith and Pat Bova at NORC; and Professor Jack P. Ruina of MIT for providing moral support and financial aid from a Carnegie Corporation grant.

NORC/N [respondents ranked likelihood that an ABM would be built from 0 to 10]:

	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*DK*
6/64	4	3	3	3	4	15	11	10	12	11	23	1
2/66	11	7	6	6	7	20	9	7	8	8	12	1

NORC/N [respondents ranked desirability of deploying an ABM from +3 to −3; response options abbreviated here]:

6/1964	*+3*	*+2*	*+1*	*0*	*−1*	*−2*	*−3*	*DK*
Around all large cities	61	17	9	2	2	1	6	1
Around your city	46	16	14	7	3	3	6	1
If risk fired by accident	29	15	15	4	7	6	9	1
If shoot down 9 of 10	67	14	8	2	1	1	4	1
If shoot down 1 of 3	42	19	17	4	4	3	4	1

NORC/N: Now I'll read a few more statements to you. Please tell me whether you strongly agree, agree, disagree, or strongly disagree with each statement, or whether you are undecided.

6/1964	*Strongly Agree*	*Agree*	*Disagree*	*Strongly Disagree*	*Undec*
If we have such missiles around our cities, we should have shelters to protect against fallout because some enemy weapons will get through the defense anyway.	18	65	7	2	8
Such missiles will make the Russians think that we are going to start a war; therefore, they might start one.	3	21	53	8	13
Antimissile missiles will lead to a stepping up of the arms race.	4	44	33	2	16
Such missiles will cost too much money to be worth while.	3	14	58	10	17
New offensive weapons will soon put antimissile missiles out of date, so it really is not worth having them.	1	12	53	7	25
Putting in antimissile missiles makes it more difficult to reach agreements on arms control and disarmament.	2	28	47	4	18
Antimissile missiles will make America stronger, and an enemy will be even less likely to attack us than without these missiles.	15	61	13	2	8

NORC/N [respondents evaluated types of cities that should be defended; q abbrev.]:

6/1964	*Most*	*Next Most*	*Least*	*Next Least*
Cities with the largest population	19	12	4	10
Cities with the most industry	24	36	3	4
Cities with military establishments around them	43	25	3	3
Cities that are important in American history	1	2	48	32
Sea and land transportation centers	9	20	7	9
Cities which are drawn by lot to insure that all have equal chances to be defended	2	2	33	39

ORC/N [respondents evaluated desirability of ABMs under certain circumstances; q abbrev.]:

6/1968	*−3*	*−2*	*−1*	*0*	*+1*	*+2*	*+3*	*DK*
How desirable is it to put ABMs around all large cities in America?	5	3	3	9	9	11	48	11
How desirable is it if they only could shoot down about one out of every two missiles?	22	7	5	9	9	10	23	16
A bigger program costing about 40 billion dollars has been suggested. How desirable is such a program to you?	14	4	5	16	9	8	26	17
If we know that the Russians have an active ABM program, how desirable would it be that the U.S. have one?	3	1	1	5	5	9	66	9

ORC/N: Here are some possible consequences of this ABM program. I would like to know for each one whether you strongly agree, agree, disagree, or strongly disagree that it might be a consequence.

6/1968	*Strongly Agree*	*Agree*	*Disagree*	*Strongly Disagree*	*DK*
Having ABMs around our cities will make people think that war is more likely	13	37	36	6	8
If we have ABMs, the Russians will think that we are going to start a war	5	25	50	9	11
Having ABMs will make America stronger and an enemy will be less likely to attack	26	48	15	3	9
New offensive weapons will put ABMs out of date, so it is really not worth having them	4	17	43	9	27

HARRIS [respondents asked to evaluate various defense options; q abbrev.]:

12/1968	*Ought to Be Done*	*Should Not Do*	*Not Sure*
Build up a system of antimissile defenses.	60	23	17
Convert the space program into a system of nuclear weapons space stations	25	49	26

AIPO [asked of those who had heard of the ABM]: Do you happen to have an opinion about the ABM program as submitted to Congress by President Nixon?

	Yes	*No*	*Not Heard*
4/69	40	27	33
5/69	41	31	28
7/69	42	27	31

AIPO [asked of those who had heard of the ABM and had an opinion]: Do you favor or oppose the ABM program submitted by Nixon?

	Favor	*Oppose*	*Not Heard or No Opinion*
4/69	25	15	60
5/69	25	14	61
7/69	23	18	59

HARRIS [prefatory description deleted]. Do you tend to approve or disapprove of President Nixon's decision to go ahead with the antimissile system?

4/1969 *Approve* (48%) *Disapprove* (25%) *Not Sure* (27%)

HARRIS [Respondents asked to agree/disagree with statements]:

4/1969		*Agree*	*Disagree*	*Not Sure*
Many scientists think the system is a mistake because it could not prevent the loss of millions of U.S. lives in an atomic attack.		40	28	32
By taking this step, President Nixon hurt the chances of reaching arms control agreements with the Russians.		19	52	29
We could have used the seven billion dollars better for our education, health, housing, and poverty needs at home.		49	31	20
7/1969				
The Safeguard ABM anti-missile system is necessary in order for us to keep up with the Russians in the nuclear arms race.		50	24	26
4/1969 and 7/1969				
It is better to be overprepared militarily than to be caught short without proper defenses.				
	4/1969	78	9	13
	7/1969	84	7	9
4/1969 and 7/1969				
In a nuclear age, the U.S. should concentrate on peaceful arms control rather than building more missile systems like the Safeguard ABM.				
	4/1969	47	29	24
	7/1969	50	24	26

Star Wars

SIND: If the U.S. had the capability of changing this situation [that we have no defense against ballistic missiles] by deploying an antiballistic missile defense, would you favor it being done?

	Yes	*No*	*No Opinion*
8/1982	86	10	4

HARRIS [respondents asked to agree/disagree with statements]:

While it might be possible to develop such defensive systems in outer space, once such means of destruction were built, they could easily be turned into threatening mankind with new and frightening space wars.

	Agree	*Disagree*	*Not Sure*
4/1983	72	24	4

Once the Russians knew we were building a new outer space antinuclear defense system, they would be more willing to negotiate a treaty that would halt the nuclear arms race.

4/1983	40	55	5

The only way to avoid a nuclear war is to develop new weapons in space that can shoot down all nuclear missiles.

	38	57	5

All in all, do you favor or oppose spending billions of dollars for the U.S. to develop a laser-beam and particle-beam outer space defense system?

4/1983	36	58	5

ABC/WP (asked of 65% who had heard/read of Reagan SDI plan): Do you favor or oppose developing such defensive weapons, or what?

	Favor	*Oppose*	*No Opinion*
4/1983	54	37	8

ABC/WP: Just your best guess, if the United States does develop such defensive weapons, would that increase or reduce the arms race between the United States and the Soviet Union, or what?

	Increase	*Decrease*	*Not Change* (vol)	*No Opinion*
4/1983	57	24	9	10

CBS/NYT: President Reagan has proposed developing a defensive system that would destroy incoming Russian missiles before they reach the United States. Some people say it might be difficult technologically, but we should try to develop it. Other people say it would be impractical, expensive and sounds like science fiction. Do you think we should try to develop the system, or not?

	Should	*Should Not*	*Don't Know*
4/1983	67	25	8

P&S: Do you favor or oppose plans to develop defensive weapons that would operate in space in order to protect the U.S. by destroying any incoming missiles?

	Favor	*Oppose*	*Don't Know*
4/1984	75	17	8

HARRIS [preliminary wording about U.S.–Soviet negotiating differences deleted here]:

Do you favor or oppose the U.S. and the Soviet Union sitting down to negotiate an agreement to outlaw the use of weapons in outer space?

	Favor	*Oppose*	*Don't Know*
7/1984	82	16	2

HARRIS: How optimistic are you that the U.S. and Russia will come to an agreement anytime soon to outlaw the use of weapons in space—very optimistic, somewhat optimistic, not very optimistic, or not optimistic at all?

	Very Optimistic	*Somewhat Optimistic*	*Not Very Optimistic*	*Not Optimistic at All*
7/1984	9	23	36	31

ABT: Would you support a U.S. attempt to build a defensive system against nuclear missiles and bombers under the following conditions:

If it did not use nuclear explosives and was only partially effective such as being able to shoot down 90% of all Soviet nuclear missiles and bombers, and letting 10% get through?

9/1984 *Yes* (57%) *No* (35%) *Don't Know* (8%)

If it used nuclear explosives and was only partially effective such as being able to shoot down 90% of all Soviet nuclear missiles and bombers, and letting 10% get through?

9/1984 *Yes* (44%) *No* (48%) *Don't Know* (8%)

If it was expensive and increased defense spending by at least 10% and increased government spending by 3%?

9/1984 *Yes* (58%) *No* (35%) *Don't Know* (7%)

CBS/NYT: Some people have proposed a defense system in space to protect the United States from nuclear attack, the concept called "Star Wars." Do you think this would make us more secure, or would it just speed up the arms race?

	Make More Secure	*Speed Up Arms Race*	*Both* (vol)	*Neither* (vol)	*No Opinion*
10/1984	31	48	3	2	16

CBS/NYT: Ronald Reagan has proposed developing a defensive nuclear system in space that would destroy incoming missiles before they reach the United States, a system some people call "Star Wars." Do you think such a system could work?

1/1985 *Yes* (62%) *No* (23%) *No Opinion* (15%)

CBS/NYT: Do you think that developing this system would make negotiations with the Soviet Union easier?

1/1985 *Yes* (48%) *No* (41%) *No Opinion* (11%)

CBS/NYT: Do you think this would be worth the amount of money it would cost?

1/1985 *Yes* (40%) *No* (46%) *No Opinion* (14%)

CBS/NYT: Would having this system make you feel secure or would putting nuclear weapons in space worry you?

1/1985 *Make Feel Secure* (25%) *Worry You* (60%) *Both* (vol. 4%)
Neither (vol. 2%) *No Opinion* (8%)

ABC/WP (asked of 63% who had heard/read of Reagan SDI proposal): Do you favor or oppose developing such defensive weapons, or what?

1/1985 *Favor* (49%) *Oppose* (44%) *No Opinion* (7%)

LAT: Generally speaking, do you favor or oppose an agreement to outlaw the use of all military weapons in outer space? Is that (favor/oppose) strongly or somewhat?

1/1985 *Favor Strongly* (43%) *Favor Somewhat* (16%)
Oppose Somewhat (15%) *Oppose Strongly* (16%)
Not Sure (10%)

AIPO: How closely have you followed the discussions over the administration's so called Star Wars proposal—that is, its proposal to develop a space-based defense against nuclear attack?

2/1985 *Very Closely* (16%) *Fairly Closely* (51%) *Not Closely at All* (30%)

AIPO: Would you like to see the United States go ahead with the development of such a system or not?

2/1985 *Yes* (52%) *No* (38%) *No Opinion* (10%)

AIPO: In your opinion, would the United States' developing this system increase or decrease the likelihood of reaching a nuclear arms agreement with the Soviet Union?

2/1985 *Increase* (47%) *Decrease* (32%) *No Difference* (vol. 13%) *No Opinion* (8%)

HARRIS [preliminary description of Reagan program deleted here]: Once the Russians knew we were building a new outer space antinuclear defense system, they would be more willing to agree to a treaty that would halt the nuclear arms race.

3/1985 *Agree* (52%) *Disagree* (44%) *Not Sure* (4%)

HARRIS: The only way to avoid a nuclear war is to develop new weapons in space that can shoot down all nuclear missiles.

3/1985 *Agree* (44%) *Disagree* (52%) *Not Sure* (4%)

HARRIS: Once it looked as though we were capable of defending against today's nuclear weapons, the Soviets would then go all-out to develop new kinds of nuclear and other weapons we couldn't defend against.

3/1985 *Agree* (75%) *Disagree* (20%) *Not Sure* (5%)

HARRIS: Even though President Reagan wants to spend only $3.8 billion at first on developing a laser defense system, the cost could rise to $25 billion to $50 billion on a system that might not even work.

3/1985 *Agree* (63%) *Disagree* (30%) *Not Sure* (7%)

HARRIS: All in all, do you favor or oppose spending billions of dollars for the U.S. to develop a laser-beam and particle-beam outer space defense system?

3/1985 *Favor* (39%) *Oppose* (56%) *Not Sure* (5%)

SIND: Some people say that in the development of any strategic defense system that could destroy incoming missiles, the Soviet Union is far ahead of the United States, while other people are saying that the United States is far ahead of the Soviet Union. What do you think?

	Soviet Union Ahead	*Soviet Union Behind*	*Both the Same*	*Not Sure*
5/1985	34	26	17	23

ABC/WP (asked of those 41% who approved SDI): Currently the U.S. and the Soviet Union have an antiballistic missile treaty that prohibits both nations from developing certain weapons. Suppose the U.S. had to violate or abandon that treaty in order to develop the space-based weapons. Would you still favor the development of those space-based weapons, or not?

	Yes	*No*	*Don't Know*
7/1985	63	32	5

M&K: The Reagan administration is now working on this program known as Star Wars. It will attempt to build a new defensive system in outer space that could shoot down nuclear missiles fired at the U.S. Since the program is currently only a research project, it is impossible to predict how complete a defense it will provide. I'm going to read a list of four possible Star Wars systems. For each one, tell me if you would strongly support, support, oppose, or strongly oppose building each type of system.

9/1985	*SS*	*S*	*O*	*SO*	*NS*
A system that was perfect and could successfully defend against all incoming nuclear weapons.	58	28	5	4	4
A system that could protect our missile sites and some population centers, but could not guarantee the safety of many of our major cities.	12	32	34	16	6
A system designed only to protect U.S. missiles, key military bases, and Washington, D.C., but not other areas.	5	16	39	34	5
A system that could provide a complete defense against long-range nuclear missiles, but cannot defend against missiles fired from submarines or bombers.	8	32	34	17	10

M&K: Now I'm going to read several statements people have made about the Star Wars proposal. Some of the statements argue in favor of Star Wars and some argue against it. Regardless of how you feel, please listen to each statement and tell me whether you strongly agree, mostly agree, mostly disagree, or strongly disagree with each one.

9/1985	*SA*	*MA*	*MD*	*SD*	*NS*
If we build the Star Wars system the Soviets will just start building more and more weapons until they can penetrate it.	37	39	13	5	6
Star Wars is the kind of bold, new step we need to lead us away from the arms race.	18	29	26	17	11
Star Wars will make nuclear weapons impotent and obsolete.	11	22	32	22	14
It would be a dangerous mistake to expand the nuclear arms race into outer space.	27	27	25	12	4
Star Wars will encourage the Soviet Union to reach an agreement with us to reduce nuclear weapons.	16	31	27	18	9
The Soviet leaders are extremely concerned about Star Wars, which means we must be on the right track with this idea.	28	36	18	8	10
Even if it's a good idea, Star Wars costs too much to seriously consider building, especially at a time when our federal deficit is already at an all-time high.	29	29	21	11	9

P&S: If a nuclear freeze were negotiated, should development of such defensive space weapons continue or should development be abandoned?

4/1984 *Continue* (54%) *Abandon* (35%) *Don't Know* (11%)

LAT: Some people say that the first thing we have to do is conduct Star Wars research to find out if the idea works. Then when we know what is at stake, there will be time enough to negotiate with the other side. But other people say that, just as soon as one side starts research, the other side will try to improve its weapons so they can avoid Star Wars defenses, and nuclear development escalates. Do you think the two sides should wait until research turns up something before they negotiate or do you think the research should be banned at the start to avoid nuclear weapons escalation?

1/1985 *Ban* (55%) *Delay* (33%) *Not Sure* (11%) *Refused* (1%)

LAT: The idea behind our nuclear defense over the past 40 years is called "Mutual Assured Destruction." It says that what keeps both the United States and the Soviet Union from firing off their nuclear weapons at each other is the certainty that the other side would counterattack with dreadful destruction. Now some people think the "Star Wars" system would destabilize this balance of terror. They claim that if one side had a way to defend itself against nuclear counterattack, it could launch an aggressive war without fear of retaliation. Do you think that a "Star Wars" system threatens its opponents or not?

1/1985 *Yes* (56%) *No* (28%) *Not Sure* (15%) *Refused* (1%)

AIPO: In your opinion, would developing this system make the world safer from nuclear destruction, or less safe?

2/1985 *Safer* (50%) *Less Safe* (32%) *No Difference* (vol 11%) *No Opinion* (7%)

ABC/WP: Have you read or heard about plans by the Reagan administration to develop weapons in outer space that could destroy nuclear missiles fired at the United States by the Soviet Union or other countries? (84% responded yes) Reagan calls the research on these weapons SDI, for Strategic Defense initiative, and some people refer to it as "Star Wars."

Supporters say such weapons could guarantee protection of the United States from nuclear attack and are worth whatever they cost. Opponents say such weapons will not work, will increase the arms race, and that the research will cost many billions of dollars. How about you: would you say you approve or disapprove of plans to develop such space-based weapons?

7/1985 *Approve* (41%) *Disapprove* (53%) *Don't Know* (5%)

M&K: Supporters of the ABM treaty say it has made the world safer by giving both the United States and the Soviet Union the security of knowing that the other side cannot use nuclear weapons without being destroyed in retaliation. At some point, development of Star Wars will require the U.S. to either withdraw from the ABM treaty or violate the treaty. If you had to choose between developing Star Wars or keeping the ABM treaty, which would you choose?

9/1985 *Develop Star Wars* (37%) *Keep ABM Treaty* (48%) *Not Sure* (15%)

AOC 10-1

POPQAE 50 (1) 143–284 (1986)
ISSN 0033-362X

Public Opinion Quarterly

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors *David Pritchard*

Interpersonal Communication and News Comprehension *John P. Robinson and Mark R. Levy*

Measuring Change in Personal Economic Well-Being *Steven J. Rosenstone, John Mark Hansen, and Donald R. Kinder*

Race Differences in Abortion Attitudes *Elaine J. Hall and Myra Marx Ferree*

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures *James B. Lemert*

Election Predictions: An Empirical Assessment *William Buchanan*

The Presence of Others and Overreporting of Voting in American National Elections *Brian D. Silver, Paul R. Abramson, and Barbara A. Anderson*

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions *George F. Bishop, Alfred J. Tuchfarber, and Robert W. Oldendick*

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys *Robert M. Groves and Lou J. Magilavy*

An Experiment with a Confidentiality Reminder in a Telephone Survey *James H. Frey*

THE POLLS: Environmental Protection *John M. Gillroy and Robert Y. Shapiro*

BOOK REVIEWS

SUMMER 1986

Published for The American Association for Public Opinion Research
by The University of Chicago Press

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

LEO BOGART BURNS W. ROPER
NORMAN M. BRADBURN BERNARD ROSHCO
ANDREW KOHUT AND ELEANOR SINGER, COUNCILLORS-AT-LARGE

Editor: ELEANOR SINGER *Managing Editor:* INA C. JAEGER
Department Editors:
PHILIP A. HARDING, *News and Notes*
MARC B. GLASSMAN, *Book Reviews*
TOM W. SMITH, *The Polls*

EDITORIAL BOARD

JAMES R. BENIGER
JOAN S. BLACK
NORMAN M. BRADBURN
CHARLES F. CANNELL
MURIEL G. CANTOR
ALBERT H. CANTRIL
DONALD DILLMAN
MARTIN R. FRANKEL
HERBERT J. GANS
ALBERT E. GOLLIN
ROBERT HORNIK
MARY R. JACKMAN
F. THOMAS JUSTER
RONALD C. KESSLER
MARK R. LEVY
PHILIP MEYER
J. RONALD MILAVSKY
WARREN J. MITOFSKY
ELISABETH NOELLE-NEUMANN
STANLEY PRESSER
DONALD F. ROBERTS
MORRIS ROSENBERG
CHARLES R. WRIGHT
DANIEL YANKELOVICH

Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research. Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 SUMMER 1986 NUMBER 2

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors *David Pritchard* 143

Interpersonal Communication and News Comprehension *John P. Robinson and Mark R. Levy* 160

Measuring Change in Personal Economic Well-Being *Steven J. Rosenstone, John Mark Hansen, and Donald R. Kinder* 176

Race Differences in Abortion Attitudes *Elaine J. Hall and Myra Marx Ferree* 193

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures *James B. Lemert* 208

Election Predictions: An Empirical Assessment *William Buchanan* 222

The Presence of Others and Overreporting of Voting in American National Elections *Brian D. Silver, Paul R. Abramson, and Barbara A. Anderson* 228

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions *George F. Bishop, Alfred J. Tuchfarber, and Robert W. Oldendick* 240

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys *Robert M. Groves and Lou J. Magilavy* 251

An Experiment with a Confidentiality Reminder in a Telephone Survey *James H. Frey* 267

THE POLLS: Environmental Protection *John M. Gillroy and Robert Y. Shapiro* 270

Book Reviews *Marc B. Glassman, Editor* 280

Sandra J. Ball-Rokeach, Milton Rokeach, and Joel W. Grube, *The Great American Values Test: Influencing Behavior and Belief Through Television* *Melvin M. Mark* 280

Willard D. Rowland, Jr. and Bruce Watkins, Eds., *Interpreting Television: Current Research Perspectives* *Muriel G. Cantor* 282

NOTICE TO CONTRIBUTORS

Three copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic-references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

Public Opinion Quarterly is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

Public Opinion Quarterly has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

Homicide and Bargained Justice: The Agenda-Setting Effect of Crime News on Prosecutors

DAVID PRITCHARD

THE EFFECT of publicity upon the adversary system of criminal justice has received an enormous amount of attention in recent years. The Supreme Court of the United States has addressed the fair trial/free press issue several times, and many social scientists have studied the effect of publicity on potential jurors (for reviews of the empirical research see Bush, 1970; Connors, 1975; Simon, 1977; Buddenbaum et al., 1981).

Generally overlooked in the fair trial/free press debate, however, is the fact that as many as 90 percent of all criminal convictions in the United States are the result of plea bargaining rather than full-blown adversary trials (Heumann, 1978; Brosi, 1979). The defendant admits guilt in return for some implicit or explicit concession from the prosecution. Because there is no jury, there is no chance that press coverage will prejudice the jury.

Nonetheless, it is possible that press coverage will taint the *process*. This study addresses that issue by examining the relationship between

---

Abstract This article reports the findings of research into the possibility that differences in newspaper coverage of individual criminal cases may influence the behavior of key justice-system officials with respect to those cases. The study analyzed police and court records regarding all people arrested for homicide over an 18-month period in Milwaukee County, Wisconsin. Newspaper coverage of those cases also was analyzed. The results suggest that the amount of space newspapers devote to a criminal case helps set the agenda of at least one class of public officials—prosecutors who must decide which criminal cases to plea-bargain, and which to take to trial.

David Pritchard is Assistant Professor at the Indiana University School of Journalism, Bloomington, Indiana 47405. An earlier version of this article was presented to the Annual Meeting of the Law and Society Association, San Diego, California, in June 1985. The author wishes to thank Professors William A. Hachten, Robert E. Drechsel, Sharon Dunwoody, and Stewart Macaulay for insightful comments on previous drafts.

Public Opinion Quarterly Vol. 50:143–159 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-143/$2.50

newspaper coverage of individual cases and whether prosecutors engage in plea bargaining in those cases.

It is an important issue to consider, because the results of the consensual process epitomized by plea bargaining are quite different from those of the adversarial process epitomized by jury trials. For example, defendants found guilty by a jury are more likely to be sentenced to a period of incarceration, and can expect to receive substantially longer sentences than those whose cases are plea bargained, everything else being equal (Uhlman and Walker, 1980; Shane-DuBow et al., 1981; Brereton and Casper, 1982; Pruitt and Wilson, 1983).

Previous Work

Prosecutors are avid readers of newspaper stories about their cases, and most say that the news media are good indicators of the public image of the criminal justice system (Drechsel, 1983). So it is not entirely surprising that some prosecutors acknowledge that they take press coverage of a case into consideration in deciding whether to engage in plea negotiations.

For example, a study of plea bargaining in two California jurisdictions found that prosecutors and defense lawyers were less likely to agree in negotiations if a case had received news coverage (Utz, 1978). Surveys of prosecutors and public defenders in Cook County, Illinois, also suggested that a substantial number of prosecutors would not plea bargain in cases that had received news coverage (Jones, 1978). One Cook County prosecutor said: "I'd be crazy to bargain in any case that received a lot of news coverage. The public will demand justice, which means that I usually go to trial" (Jones, 1978:200). A Cook County public defender claimed that prosecutors refused to negotiate in publicized cases because they are "afraid of being considered soft on criminals" (Jones, 1978:202; see also Kaplan, 1965; Miller, 1970; and Eisenstein and Jacob, 1977).

Why might publicity make a prosecutor unwilling to plea bargain a case? Alschuler (1968) studied prosecutors in a dozen large American cities, and found that they were motivated more by what they perceived to be their self-interest than by considerations of justice or fairness for the defendant.

> It is politically embarrassing for a prosecutor to consent to a light sentence for a serious offender, even as part of a plea agreement. It is usually easier, in fact, for a prosecutor to explain an acquittal than it is for him to explain a plea agreement in a publicized case; an acquittal can be attributed to the finder of fact. In routine cases, by contrast, there is little danger that sentencing will

> attract public attention. A prosecutor can maximize the political value of these cases by securing the maximum number of convictions, regardless of the price he pays in sentencing concessions (Alschuler, 1968:107).

In other words, prosecution is a political process, and prosecutors have a political stake in how their actions are perceived. Maintaining a public image as a crime-fighter is important to the prosecutor, perhaps to the extent of stressing adversary dispositions in publicized cases, regardless of the strength of evidence against the defendant. That way, if something "good" (like a conviction) happens, the prosecutor can take the credit. If something "bad" (like an acquittal or dismissal of charges) happens, the prosecutor can implicitly or explicitly shift the responsibility to the judge or to the jury. The result is that the blame is transferred (Newman, 1966; Galanter, 1983). Prosecutors may be especially likely to act this way in homicide cases (Alschuler, 1968).

In some cases, however, prosecutorial self-interest can lead to negotiations. This is most likely to happen when the prosecutor feels a need to get a conviction—any kind of conviction—despite weak evidence against the defendant. Alschuler writes: "Political considerations may, on occasion, make it important for a prosecutor to secure a conviction for a particular crime, and plea negotiations may provide the only practical means of achieving this objective" (Alschuler, 1968:109).

However, prosecutors' most common reaction to publicity, the existing research makes clear, is a desire to avoid being perceived as soft on criminals.

The notion that the press may influence prosecutorial decision making has clear parallels with the familiar agenda-setting hypothesis, which suggests that the relative prominence of an issue in the news media will influence how salient that issue is to members of the audience (McCombs and Shaw, 1972; McLeod et al., 1974; Becker et al., 1975; Shaw and McCombs, 1977).

Most agenda-setting research has focused on possible effects on ordinary citizens. Only a few studies have explicitly tested the hypothesis that the press may help set the agendas of public officials. Gormley (1975) found mixed evidence for an agenda-setting effect of newspapers on state legislators. Lambeth (1978) concluded that the press helped set energy policy makers' agendas. In a study notable for the fact that it used a direct, rather than a self-reported measure of public-official behavior, Gilberg et al. (1980) found that press content set the agenda for Jimmy Carter's 1978 State of the Union speech. Swank et al. (1982) suggested, but did not test, the hypothesis that newspapers contribute to the salience of crime on local political agendas. Cook et al. (1983) found that a televised investigative report changed the agenda not only of citizens but also of elected officials.

On the other hand, Walker (1977) concluded that for three safety-related issues, the agenda of the U.S. Senate set the agenda for the *New York Times*. And Protess et al. (1985) found that the only agenda affected by a newspaper investigative series was future press coverage of the subject of the series.

In addition to the agenda-setting literature, research into reporter-source interactions contains considerable speculation (and some anecdotal evidence) suggesting that journalists' decisions about which stories to play up can influence the behavior of public officials (see, e.g., Matthews, 1960; Cohen, 1963; Dunn, 1969; Sigal, 1973; Weiss, 1974; Miller, 1978; and Peters, 1980).

Despite these occasional forays, how the news media affect the behavior of public officials is an "inadequately researched area" (Rivers et al., 1975:223). Rivers et al. suggested that research should attempt to determine under what conditions publicity alters "personal and departmental resources in ways that affect the decision-making process" (1975:231).

Such research is especially important in the criminal justice area. Stanga noted that "the influence the press has, if any, on the prosecutor is a neglected area of inquiry" (1971:264). More recently, Buddenbaum et al. came to a similar conclusion, suggesting that "a study of the effects of pretrial publicity on the disposition of cases might also be useful, to establish how much influence pretrial publicity has on the way cases are handled and on the amount of time and money spent for similar cases that receive varying publicity" (1981:7).

The Context of the Study

This study focuses on prosecutors in the district attorney's office in Milwaukee County, Wisconsin. The basic business of prosecutorial organizations is prosecuting criminal cases. In Milwaukee, as elsewhere in the United States, such cases are processed primarily by plea bargaining, which serves the cause of conserving relatively scarce organizational resources (Brosi, 1979). At the same time, however, it is not a popular way of settling cases (Hearst Corporation, 1983). In Milwaukee, for example, less than a quarter of the population favors plea bargaining (Metropolitan Milwaukee Criminal Justice Council, 1980). In addition, one of the city's newspapers flatly opposes the practice (Wills, 1977), while the other is merely skeptical (Milwaukee Journal, 1981).

Although in theory defendants decide whether to plead guilty or to exercise their constitutional right to a jury trial, in practice the prosecutor generally controls whether a case is plea bargained (Blumberg, 1967; Alschuler, 1968; Casper, 1972; Alschuler, 1975; Heumann, 1978; Gifford, 1981; Gifford, 1983). On the one hand, limited resources compel prosecutors to use plea bargaining as the normal mode of settling criminal cases. On the other hand, the public's anger about crime suggests that the electorate may not respond warmly to a prosecutor who plea bargains a visible case.

Clearly, then, not all cases can go to full-blown adversarial trial (resource constraints), but some trials are necessary (the influence of public opinion).

This study hypothesizes that newspaper coverage of a criminal case influences whether prosecutors engage in plea bargaining in a given case, and that the more extensive the newspaper coverage of a case, the less likely the district attorney's office is to negotiate in the case.

Data

Data to test this study's hypothesis were extracted from police and court records and from news stories. Information was obtained on every nonvehicular homicide case presented to the district attorney's office for possible prosecution during the 18-month period between January 1, 1981 and June 30, 1982. The study focuses on homicides because lesser crimes seldom receive press coverage in a major metropolitan community like Milwaukee. Three homicides that had not been disposed of by May 31, 1983 were eliminated from the analysis. Prosecutors filed homicide charges against every suspect arrested for homicide during the study period. In all, the cases of 90 homicide defendants were included in this study.

To find out how the cases were processed, every publicly available document on each case was scrutinized. Included were the inmate registration log at the Milwaukee County Jail; case files at the office of the Milwaukee County Clerk of Courts, Felony Division, which contained copies of criminal complaints, autopsy reports and other pieces of documentary evidence, summaries (and sometimes transcripts) of hearings, and memos from the prosecution and defense; and all news items about the cases published by Milwaukee's daily newspapers, both owned by the Journal Company, which granted full access to its files of clippings. The editions of the morning *Sentinel* and the evening and Sunday *Journal* that circulate in Milwaukee County carried 744 staff-written items—news stories, editorials, and columns of staff opinion—about the homicides in this study.

Measures

The independent variable in this study is the level of newspaper interest in the defendant's case. The dependent variable is the behavior of the prosecutor's office with respect to the case.

LEVEL OF NEWSPAPER INTEREST IN A CASE

Conceptually, this variable is a function of the resources the newspaper organization is willing to devote to a case. Although newspapers have at their disposal several kinds of resources, the two that are most visible to outsiders like prosecutors are space in the paper and staff effort.

The amount of space a newspaper devotes to a case can be measured in a straightforward manner. The number of news items about a case is an interval-level variable. So is the number of paragraphs written about a case. This study divides total number of paragraphs by total number of stories to come up with a measure of the level of newspaper interest in the case: the average length of news items about the case, excluding editorials (which tend to be quite short) and excluding stories published after the prosecutor's discretionary period.

That discretionary period is defined as lasting from the date of the homicide until the mode of disposition was known, generally at the beginning of a trial or the acceptance of a guilty plea by a judge. Ending the discretionary period any earlier would skew the results of the analysis, because plea agreements can be—and in several cases were—made or unmade at the last minute. Cases that seem irrevocably headed toward adversary dispositions sometimes are negotiated on the eve of a scheduled trial. Similarly, defendants may accept plea bargains, but then change their minds at the hearing and refuse to plead guilty.

Average story length, rather than total number of stories or paragraphs, is used because average story length in theory is not a function of how long a case lingers in the felony disposition process. The total number of stories or paragraphs the press devotes to a case, on the other hand, can be more a function of how many pretrial hearings take place than of how interested the press is in the case.

Number of paragraphs, rather than number of column inches, is used to measure story length because varying column sizes in both newspapers made computing standardized column inches very difficult. Standardized column inches and number of paragraphs are highly correlated, however (Budd, 1964).

Both Milwaukee newspapers covered homicides similarly (Pritchard, 1985). For that reason, this study combines the newspapers' coverage to form a single variable: the average length of the stories the Milwaukee newspapers published about a defendant's case. Separate analyses were conducted for each newspaper's coverage, with results virtually identical to the result produced by the combined coverage variable.

The newspapers published no stories at all about 5 of the 90 homicide prosecutions in this study. Those cases are coded as having an average story length of 0.

The amount of newspaper staff effort devoted to a given case was measured by analyzing newspaper content. The typical piece of crime news comes directly from routine law enforcement or judicial sources, often documentary sources such as police blotters, jail logs, and criminal complaints (Stanga, 1971; Cohen, 1975; Sherizen, 1978; Drechsel et al., 1980). Reporters who cover police and court beats can gather such news with relatively little effort.

On some cases, however, reporters do more. They may use nonroutine sources such as witnesses to the crime, friends and relatives of the suspect and/or victim, or not-for-attribution comments from law enforcement officials. Editorials and staff-written columns of opinion also represent nonroutine kinds of coverage.

Accordingly, stories about homicide cases can be categorized either as "routine" (if only routine sources were used) or "nonroutine" (if at least one nonroutine source was used). Of the 744 news items in this study, 13.4 percent were nonroutine by this definition.[1] The variable created using the routine/nonroutine dichotomy is the proportion of nonroutine stories in the newspapers' coverage of a case.

This study did not attempt to determine whether predisposition coverage of individual cases was fair or accurate, however those terms might be defined. The independent variables attempt only to operationalize the level of press interest in a case, measured by the amount of organizational resources newspapers devoted to covering the case.

PROSECUTOR'S PLEA BARGAINING BEHAVIOR

Plea bargaining in this study is measured not by whether a case was settled consensually, but by whether the prosecutor actually engaged in plea negotiations.This "negotiated/did-not-negotiate" variable measures the prosecutor's actual behavior, not whether that behavior led to a consensual outcome. In theory, the negotiated/did-not-negotiate variable is untainted by the behavior of other actors in the criminal justice system. For example, although negotiations cannot succeed without the participation of the prosecutor, a willing prosecutor can be stymied by a defense lawyer or defendant unwilling to accept the prosecutor's terms. Such cases would be disposed of adversarially, but not because the prosecutor failed to negotiate.

---

[1] All coding was done by the author. To test intercoder reliability, a colleague coded a 10 percent sample of the 744 news items. There was agreement on 96 percent of the items, resulting in a reliability coefficient (Scott's pi) of .85 (Scott, 1955).

Court records and/or news items contained explicit evidence that prosecutors engaged in negotiations in 45 of the 90 cases in this study. In 35 of the cases in which prosecutors negotiated, the result was a consensual settlement, in which the prosecutor and the defense agreed on the appropriate disposition of the case. In the remaining 10 cases, defendants refused plea bargains offered by prosecutors. Those cases either went to jury trial or were dismissed over the objection of the prosecution.

CONTROL VARIABLES

One of the most challenging tasks in a study of this nature is to hold constant factors that may affect both the independent and the dependent variables. In an attempt to achieve that goal, this study uses an extensive set of control variables. Much of the variation in crime seriousness, a factor that could influence both newspaper interest and prosecutors' plea bargaining behavior, is implicitly controlled by this study's exclusive focus on homicides.

Homicides themselves can vary in a number of ways, however, so additional control variables are used. They include personal attributes (race, age, and sex) of the homicide suspect and victim; whether the suspect and the victim knew each other; the suspect's prior record; the initial charge against the defendant (first-degree murder or a lesser homicide charge, such as manslaughter); whether the defendant was charged with crimes beyond the first homicide count; and the number of suspects alleged to have been involved in the homicide.

Some factors, such as the prominence of the homicide suspect and victim or the bizarreness of the incident, are difficult to quantify. In most cases, however, court records and news stories contain enough details of the incident and of the people involved in it to permit a qualitative evaluation of such aspects of the case.

Results

Discriminant analysis (Cohen and Cohen, 1975; Klecka, 1975) was used to test the hypothesis concerning the influence of newspaper coverage on plea bargaining behavior. The analysis supported the hy-

Table 1. Results of Discriminant Analysis, with Whether the Prosecution Negotiated as the Dependent Variable

Canonical correlation squared	.238
Improvement in ability to predict	53.3%
Relative contribution of significant discriminating variables:	
Average story length	34.8%
Suspect knew victim	26.9%
Prior record	19.5%
Initial charge	12.4%
Multiple charges	6.5%

pothesis: press behavior—specifically, the average length of stories about a case—was the strongest predictor of whether prosecutors engaged in negotiations. The proportion of stories about a case that relied partly or entirely on nonroutine sources, however, was not a significant predictor of negotiations.

Table 1 shows the results of the analysis. Of all the variables, five proved to be statistically significant (at $p = .05$) predictors: average story length (the shorter the average story, the more likely the prosecutor would negotiate); whether the defendant and the victim knew each other (negotiations were more likely if the defendant and victim had been acquainted); the defendant's prior record (negotiations were more likely if the defendant had no prior record); the seriousness of the initial charge (negotiations were more likely when charges were less serious);[4] and whether the defendant faced multiple charges stemming from the incident (negotiations were more likely when there was only one charge).

The canonical correlation for the discriminant function is .4883, which means that the variables in the function account for 23.8 percent of the total variance in whether the prosecution engaged in plea negotiations. Average story length is the strongest predictor of prosecutorial behavior, contributing more than a third of the variance accounted for by the function, 34.8 percent, which is 8.3 percent of the entire variance.

Easier to understand, perhaps, is the fact that the discriminant function correctly predicts the prosecutor's negotiating behavior in 69 of the 90 cases (76.7 percent). Without the information contributed by the variables in the function, successful predictions could be made only half of the time. In other words, knowing the values of the variables in the function provides a 53.3 percent improvement in predictive ability over chance guessing. The function is statistically significant at the .0003 level.

Despite the statistically significant results of the quantitative analysis, the cases in which the discriminant function makes an incorrect prediction merit some attention. In 5 of those 21 cases, the discriminant function computes a case's probability of membership in a given category (negotiated or not negotiated) at more than 80 percent.[2]

There was no explicit evidence of negotiations in two of these extreme outlier cases (*Wisconsin* v. *Reynosa, Wisconsin* v. *Shelton*), despite circumstances in both—minimal press attention, previous relationship with victim, no prior criminal record, etc.—that strongly predict bargaining. Both cases went to trial, and both defendants were convicted.

The striking fact about the cases, however, is that both defendants may sincerely have believed they were innocent. Virginia Reynosa, for example, did not actually stab the person whose death led to her first-degree murder conviction; she merely urged her son on as he wielded the knife. In addition, there was some evidence of provocation on the part of the victim. The son was convicted of murder for the stabbing; the mother was convicted under a portion of Wisconsin law that makes parties to a crime as guilty as the actual perpetrators. The charges are legally equivalent, however, and in Wisconsin a first-degree murder conviction carries a mandatory life prison sentence.

Jessie Shelton said that she had been the victim of domestic violence, and shot her husband in self-defense. The prosecutor who took her to trial for manslaughter and the jury that convicted her may have been skeptical of that claim; Jessie Shelton weighed 290 pounds, and her husband was blind. The judge, however, was more sympathetic. Saying he found an element of self-defense in the shooting, he placed Shelton on five years' probation rather than sending her to prison.

Incidentally, Shelton's husband was the most prominent defendant or victim in this study. He owned a taxi company serving Milwaukee's black community, and he had been active in community affairs. A measure of his prominence is that he was the only defendant or victim who had been the subject of non-crime-related news coverage before the homicide. Despite that—and despite Jessie Shelton's potentially newsworthy claim that she killed her husband in self-defense—pretrial press coverage of the case was minimal (a total of three relatively short routine stories.)

At least from a social point of view, Reynosa and Shelton may have believed, rightly or wrongly, that they were innocent. If so, it is not surprising they would be unlikely to be receptive to negotiations that would lead to guilty pleas and prison terms, as almost always is the case when a homicide case is plea bargained.

---

[2] It should be noted, of course, that in 19 other cases the discriminant function's 80 percent-or-better predictions were correct.

In the other three extreme cases (*Wisconsin* v. *Fraser*, *Wisconsin* v. *Crosley*, and *Wisconsin* v. *Murray*), prosecutors engaged in negotiations despite circumstances—intense press attention, extensive prior criminal record, multiple charges stemming from the incident, etc.—that strongly predict no negotiations. All three homicides were fairly routine. The press attention to the cases resulted from chance occurrences unrelated to the nature of the crime or of the people involved in it.

The Fraser and Crosley cases were fairly typical inner-city homicides; both began with holdups at all-night gas stations and ended with the shooting of the gas station attendants. In each case, the prosecution bargained with the defendant because it needed his testimony against an accomplice.

Those evidence-related considerations apparently outweighed the influence of extensive press coverage in the cases. But the press focused on the cases for reasons having little to do with the crimes themselves. Fraser's lawyer, for example, asked the judge to bar the press and public from the courtroom. The request, a relatively unusual one in Wisconsin, and the judge's denial of it inspired some press attention. The newspapers also singled out the Fraser case for exceptional feature coverage. The robbery and shooting took place around midnight Saturday, too late for more than a one-paragraph story in the final edition of Sunday morning's *Journal*. That gave reporters all day Sunday to interview nonroutine sources for substantial stories that would be published on Monday, typically a "slow" news day because governmental beats rarely generate news on Sunday.

In addition, five months after the incident, the *Journal* interviewed friends and family of the victim, and published a 51-paragraph Sunday feature titled "Last Day of Work—and Life." Before the suspects were arrested, the victim's family received a series of terrifying telephone calls. The *Journal* ran an 18-paragraph story ("Chilling Calls Taunted Victim's Kin").

In short, Fraser's crime was fairly routine; the coverage of it was not. It is a case that may have been plea bargained only because prosecutors had evidence problems that forced them to seek Fraser's agreement to testify against his accomplice.

The prosecution offered a similar bargain to Sylvania Crosley, charged with first-degree murder and armed robbery in connection with a crime similar to Fraser's. Crosley's crime took place in 1976, but no arrests were made until 1981, when an informant told police who had done it. By that time Crosley was in prison on unrelated charges; without his testimony, his alleged accomplice would go free.

Crosley initially agreed to the deal. But when it was time for him to testify, he refused and charges against his alleged accomplice in the gas-station holdup had to be dismissed. Defendants who renege on deals make prosecutors and judges unhappy. Crosley was convicted at trial of first-degree murder and armed robbery, and the judge imposed the maximum sentence—life plus 20 years.

Crosley's crime, like Fraser's, was fairly routine. The press coverage was routine, too, until allegations were made that money had been paid to Crosley's lawyer to buy Crosley's silence. The possibility of corruption sparked a 50-paragraph *Sentinel* story, "Probe of Payment to Lawyer Urged," and a 15-paragraph story a week later noting that authorities were probing ("Allegations on Payment of Lawyer's Fees Probed"). No other stories about the case were as many as 10 paragraphs long.

The final case that was negotiated despite intense press coverage stemmed from a tavern shooting that was routine in all but one respect—a group of Milwaukee *Sentinel* staffers was at the bar when the shooting started. Two of the journalists were seriously injured by stray bullets.

Because of weak and contradictory evidence, no arrests were made for more than a year after the shooting. Before any arrests were made, the *Sentinel* gave the case extensive publicity, including two lengthy first-person stories by one of the journalists who had been wounded. Other unusual stories during that period included a 63-paragraph story detailing victims' and witnesses' disappointment with the police department's failure to make arrests in the case and a 33-paragraph feature about the widow of the man who was killed in the shooting.

Fifteen months after the homicide, two suspects were arrested. The evidence was so weak that charges against one were dismissed outright. The other defendant, Regan Murray, faced first-degree murder and four nonhomicide charges, but agreed to plead guilty to reduced charges.

Why were the charges reduced when publicity had been so extensive? The available evidence gives no definite answer to that question, but it appears likely that the prosecution realized that the shaky evidence made conviction at trial problematic. The public pressure from the *Sentinel* made some kind of conviction imperative: thus the negotiations (Alschuler, 1968).

Discussion

This study's findings suggest that newspapers help set the plea bargaining agendas of Milwaukee prosecutors, at least in homicide cases. The amount of space newspapers were willing to devote to the typical story about a case was a stronger predictor than any other variable in this study of whether the prosecutor would negotiate.

This research found evidence suggesting an agenda-setting effect of ordinary crime news, the kind produced day in and day out by beat reporters. Recent studies from Northwestern University's Center for Urban Affairs and Policy Research, on the other hand, have examined possible agenda-setting effects only of investigative news reports (Cook et al., 1983; Protess et al., 1985).

In addition, this study did not use self-reported data. Most other research in this area has relied on public officials' own estimations of whether they have been influenced by the media (Gormley, 1975; Lambeth, 1978; Cook et al., 1983; Protess et al., 1985). Virtually all the measures of prosecutorial behavior used in this study, however, came directly from court records. The data came from cases that were already completed, ensuring that the participants were not aware that their behavior would become part of a research project. In addition, the data were recorded by court employees as part of the normal routine of the criminal courts, with no hint that the data would provide the raw material for an academic study. Unless the raw data contain a systematic bias that is not immediately apparent, they are unobtrusive measures (Webb, et al., 1966), unlike the self-reported data used as dependent measures in much of the research into the possibility that the news media may set the agenda for public officials.

That said, it must be acknowledged that documentary records are far from perfect. The facts of certain cases, such as the Reynosa and Shelton cases, strongly imply that the prosecution offered a deal that was turned down, despite the fact that court records and newspaper coverage contain no explicit evidence of negotiations.

Similarly, the examination of the five "outlier" cases revealed that the strength of evidence against a defendant (or a defendant's accomplice) could influence whether the prosecution negotiated a case. Evidence strength was not included as a quantitative variable in this study because it is extremely difficult to measure in any systematic way (Eisenstein and Jacob, 1977:182–183). In addition, evidence strength may well be a socially defined construct greatly dependent upon attributes of the relationships between the prosecutor and the defense lawyer, and between the defense lawyer and the defendant.

Finally, it is difficult to predict what effect a strength-of-evidence variable would have on plea bargaining. If evidence is weak, the prosecutor has an incentive not only to negotiate but to offer a good deal. If the evidence is strong, however, the defendant has an incentive to plead guilty even if not offered a good deal to avoid the so-called trial penalty (Uhlman and Walker, 1980; Shane-DuBow et al., 1981; Brereton and Casper, 1982; Pruitt and Wilson, 1983). Guilty pleas in cases where there is no evidence of negotiations often are implicit plea bargains (Heumann, 1978).

Related to the issue of evidence strength are questions involving the content of negotiations. Do prosecutors offer the same bargains to all similarly situated defendants, or does the defendant's race, how his or her case was covered, or other factors influence the content of the actual negotiations? This study did not have access to information about the offers and counter-offers that are a normal part of negotiations, but researchers should attempt to gain access to such information (see, e.g. Maynard, 1984).

Future research should also examine in detail the extent to which prosecutors (or other law-enforcement personnel) influence newspapers' decisions about which cases to cover and how to cover them (see, e.g., Drechsel, 1983). This study found no evidence of such an influence, but then court records and news coverage are not the best places to look for such an effect. Researchers need to get into newsrooms, prosecutors' offices, courthouse hallways, and courtrooms to find the answers to such questions.

This study focused on a narrow class of public officials—prosecutors handling homicide cases—in one community. As such, it offers no conclusions about whether the press sets the agenda for public officials in other situations and other places. Ideally, future studies will address such issues.

References

Alschuler, A.
1968 "The prosecutor's role in plea bargaining." University of Chicago Law Review 36:50–112.
1975 "The defense attorney's role in plea bargaining." Yale Law Journal 84:1179–1314.

Becker, L. B., M. E. McCombs, and J. M. McLeod
1975 "The development of political cognitions." In S. H. Chaffee, ed., Political Communication: Issues and Strategies for Research. Beverly Hills, Calif.: Sage Publications.

Blumberg, A.
1967 Criminal Justice. Chicago: Quadrangle Books.

Brereton, D. and J. D. Casper
1982 "Does it pay to plead guilty? Differential sentencing and the functioning of criminal courts." Law & Society Review 16:45–70.

Brosi, K. B.
1979 A Cross-City Comparison of Felony Case Processing. Washington, D.C.: U.S. Department of Justice.

Budd, R.
1964 "Attention score: a device for measuring news 'play'." Journalism Quarterly 41:259–262.

Buddenbaum, J. M., D. H. Weaver, R. L. Holsinger, and C. J. Brown
1981 Pretrial Publicity and Juries: A Review of Research. Bloomington, Ind.: Indiana University School of Journalism.
Bush, C. R.
1970 Free Press and Fair Trial: Some Dimensions of the Problem. Athens, Ga.: University of Georgia Press.
Casper, J.
1972 American Criminal Justice: The Defendant's Perspective. Englewood Cliffs, N.J.: Prentice-Hall.
Cohen, B. C.
1963 The Press and Foreign Policy. Princeton, N.J.: Princeton University Press.
Cohen, J., and P. Cohen
1975 Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. New York: John Wiley & Sons.
Cohen, S.
1975 "A comparison of crime coverage in Detroit and Atlanta newspapers." Journalism Quarterly 52:726–730.
Connors, M. M.
1975 "Prejudicial publicity: an assessment." Journalism Monographs 41.
Cook, F. L., et al.
1983 "Media and agenda setting: effects on the public, interest group leaders, policy makers, and policy." Public Opinion Quarterly 47:16–35.
Drechsel, R. E.
1983 News Making in the Trial Courts. New York: Longman.
Drechsel, R. E., K. Netteburg, and B. Aborisade
1980 "Community size and newspaper reporting of local courts." Journalism Quarterly 57:71–78.
Dunn, D.
1969 Public Officials and the Press. Reading, Mass.: Addison-Wesley.
Eisenstein, J., and H. Jacob
1977 Felony Justice: An Organizational Analysis of Criminal Courts. Boston: Little, Brown and Co.
Galanter, M.
1983 "Reading the landscape of disputes: what we know and don't know (and think we know) about our allegedly contentious and litigious society." UCLA Law Review 31:4–71.
Gifford, D. G.
1981 "Equal protection and the prosecutor's charging decision: enforcing an ideal." George Washington Law Review 49:659–719.
1983 "Meaningful reform of plea bargaining: control of prosecutorial discretion." University of Illinois Law Review 1983:37–98.
Gilberg, S., C. Eyal, M. McCombs, and D. Nicholas
1980 "The State of the Union Address and the press agenda." Journalism Quarterly 57:584–588.
Gormley, W. T., Jr.
1975 "Newspaper agendas and political elites." Journalism Quarterly 52:304–308.
Hearst Corporation
1983 The American Public, the Media & the Judicial System: A National Survey on Public Awareness and Personal Experience. New York: The Hearst Corporation.
Heumann, M.
1978 Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys. Chicago: University of Chicago Press.
Jones, J. B.
1978 "Prosecutors and the disposition of criminal cases: an analysis of plea bargaining rates." Journal of Criminal Law & Criminology 69:402–412.

Kaplan, J.
1965 "The prosecutorial discretion—a comment." Northwestern University Law Review 60:174–193.

Klecka, W. R.
1975 "Discriminant analysis." In Nie, N., et al. Statistical Package for the Social Sciences, second ed. New York: McGraw-Hill.

Lambeth, E. B.
1978 "Perceived influence of the press on energy policy making." Journalism Quarterly 55:11–18, 72.

Matthews, D. R.
1960 U.S. Senators and Their World. Chapel Hill: University of North Carolina Press.

Maynard, D. W.
1984 "The structure of discourse in misdemeanor plea bargaining." Law & Society Review 18:75–104.

McCombs, M., and D. L. Shaw
1972 "The agenda-setting function of mass media." Public Opinion Quarterly 36:176–187.

McLeod, J. M., L. B. Becker, and J. E. Byrnes
1974 "Another look at the agenda-setting function of the press." Communication Research 1:131–166.

Metropolitan Milwaukee Criminal Justice Council
1980 Public Opinion Survey 600.

Miller, F. W.
1970 Prosecution: The Decision to Charge a Suspect with a Crime. Boston: Little, Brown and Co.

Miller, S.
1978 "Reporters and congressmen: living in symbiosis." Journalism Monographs 53.

Milwaukee Journal
1981 "A deeper look at plea bargains." Editorial, January 11.

Newman, D. J.
1966 Conviction: The Determination of Guilt or Innocence Without Trial. Boston: Little, Brown and Co.

Peters, C.
1980 How Washington Really Works. Reading, Mass.: Addison-Wesley.

Pritchard, D.
1985 "Race, homicide and newspapers." Journalism Quarterly 62:500–7.

Protess, D. L., D. R. Leff, S. C. Brooks, and M. T. Gordon
1985 "Uncovering rape: the watchdog press and the limits of agenda setting." Public Opinion Quarterly 49:19–37.

Pruitt, C. R., and J. Q. Wilson
1983 "A longitudinal study of the effect of race on sentencing." Law & Society Review 17:613–635.

Rivers, W. L., S. Miller, and O. Grady
1975 "Government and the media." In S. H. Chaffee, ed., Political Communication: Issues and Strategies for Research. Beverly Hills, Calif.: Sage Publications.

Scott, W. A.
1955 "Reliability of content analysis: the case of nominal scale coding." Public Opinion Quarterly 19:321–325.

Shane-DuBow, S., et al.
1981 Wisconsin Felony Sentencing Guidelines: Phase I of Research and Development. Madison: Wisconsin Center for Public Policy.

Shaw, D. L., and M. E. McCombs
1977 The Emergence of American Political Issues: The Agenda-Setting Function of the Press. St. Paul: West Publishing.

Sherizen, S.
1978 "Social creation of crime news: all the news fitted to print." In C. Winick, ed., Deviance and Mass Media. Beverly Hills, Calif.: Sage Publications.
Sigal, L. V.
1973 Reporters and Officials: The Organization and Politics of Newsmaking. Lexington, Mass.: D. C. Heath.
Simon, R.
1977 "Does the court's decision in Nebraska Press Association fit the research evidence on the impact on jurors of news coverage?" Stanford Law Review 29:515–528.
Stanga, J. E.
1971 "The press and the criminal defendant: newsmen and criminal justice in three Wisconsin cities." Unpublished Ph.D. dissertation, University of Wisconsin.
Swank, D. H., H. Jacob, and J. Moran
1982 "Newspaper attentiveness to crime." In H. Jacob and R. L. Lineberry, eds., Governmental Responses to Crime: Crime on Urban Agendas. Washington, D.C.: National Institute of Justice.
Uhlman, T. M., and N. D. Walker
1980 " 'He takes some of my time, I take some of his': an analysis of sentencing patterns in jury cases." Law & Society Review 14:323–341.
Utz, P. J.
1976 Settling the Facts: Discretion and Negotiation in Criminal Court. Lexington, Mass.: Lexington Books.
Walker, J.
1977 "Setting the agenda in the U.S. Senate: a theory of problem selection." British Journal of Political Science 7:423–445.
Webb, E. J., D. T. Campbell, R. D. Schwartz, and L. Sechrest
1966 Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago: Rand McNally & Co.
Weiss, C. H.
1974 "What America's leaders read." Public Opinion Quarterly 38:1–22.
Wills, R. H.
1977 Milwaukee Sentinel Editorial Policy.

Interpersonal Communication and News Comprehension

JOHN P. ROBINSON AND MARK R. LEVY

THE PERENNIAL survey finding that the American public obtains most of its information from television is widely cited and generally accepted. Indeed, almost two-thirds of all Americans now identify television as their "main source" of news and that proportion has grown steadily since the late 1950s (Roper, 1983, 1985). While numerous researchers have raised methodological cautions about this generalization (Carter and Greenberg, 1963; Stone, 1969–70; Robinson, 1971; Stevenson and White, 1980), the "debate" over where the public gets most of its news by and large has been conducted in a conceptual and theoretical vacuum.

First of all, "main source" studies are almost always based on self-reports of information gain. Such subjective measures indicate only where people *think* they find out the news. However, self-reports of

---

Abstract This study examines the role of interpersonal and mass media channels in public awareness and comprehension of major news stories. Unlike previous research which asks about respondent perceptions of their "main source" for news, this study attempts to determine *actual* comprehension of stories that had been in the news during the previous week. On the basis of two separate probability samples, respondent awareness and comprehension of a week's news was measured and related to demographic, news media use, and interpersonal discussions variables. A Multiple Classification Analysis of the data indicates that conversation about the news is a major and often overlooked correlate of comprehension, and that interpersonal channels may play at least as important a role in the public's awareness and understanding of the news as exposure to the news media.

John P. Robinson is the Director of the Survey Research Center and a Professor of Sociology at the University of Maryland. Mark R. Levy is an Associate Professor of Journalism and a Research Associate of the Center for Research in Public Communication, University of Maryland. An earlier version of this article was presented to the 40th annual meeting of the American Association for Public Opinion Research, McAfee, NJ, May 1983. Support for this study was provided by the John and Mary Marke Foundation.

Public Opinion Quarterly Vol. 50:160–175 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-160/$2.50

channel influence measure neither *actual* information gain nor, perhaps more important, how well that information is comprehended. These shortcomings have hindered our ability to understand the role of the mass media in the opinion formation process and suggest the need for studies based on better measures of information gain and comprehension.

Second, when asked "main source" questions, fewer than 1 respondent in 20 typically points to interpersonal channels as their principal news source. This finding stands in marked contrast to several well-established paradigms of public opinion and mass communication research. For example, the importance of "people" as an information source has been demonstrated in numerous studies of political communication effects (e.g., Katz and Lazarsfeld, 1955; Tichenor et al., 1973; Robinson, 1976), diffusion of innovations (e.g, Mason, 1963; Rogers, 1983), and communication about "major" news stories (e.g., Troldahl and Van Dam, 1965; Greenberg, 1964; Gantz, 1983). On the basis of this literature, then, one might expect to find previously overlooked linkages between interpersonal communication, information gain, and news comprehension.

The study presented here examines the comprehension of news and relates levels of comprehension to the use of mass media and interpersonal communication channels by individuals. This study is informed by the revision of the two-step flow hypothesis outlined in Robinson (1976) and by the general literature on interpersonal communication (see, for example, Schramm, 1973; Chaffee, 1982). It assumes that information is often, but not always, characterized by a "horizontal" flow between reasonably well-informed and interested individuals, and that information exchange in interpersonal settings is facilitated by such factors as greater personal involvement, immediate feedback, tailoring of messages, and the like.

We further assume that interpersonal discussions may be an important and largely overlooked agent in what Gerbner et al. (1980) have hypothesized is a "mainstreaming" process. Given that much of what individuals know about "the world out there" comes to them from the mass media *and* through interpersonal channels, the interaction of media messages and interpersonal discussion becomes central in considering the mechanisms that create possible "mainstreaming" effects. It is, after all, well documented that interpersonal discussions sometimes exercise a social control function, pressuring individuals into conformity with generally accepted perceptions and understandings. If those attitudes, perceptions, and understandings originate even partly in media messages, then talking about the news may simultaneously produce audience understandings which are widely shared and which are highly similar to journalistically created accounts of social reality.

Method

Unlike previous studies of news comprehension which have experimentally manipulated message attributes (e.g., Findahl and Höijer, 1976; Katz et al., 1977; Gunter, 1980) or which have sought to determine recall and comprehension of a single evening's news (e.g., Neuman, 1976; Robinson et al., 1982), this study focuses on actual national and international news messages, published and broadcast over a week's time.

While there is nothing magical about a seven-day period, asking about news at the end of a week may more realistically reflect actual patterns of audience exposure. The chance, for instance, that an individual will have watched any given evening's network news may be as small as 1 in 20 (Robinson, 1971; Lichty, 1982). However, over the course of a week, the likelihood of any exposure to network news and its associated opportunity for information gain and comprehension rises substantially.

Additionally, over the course of a week, there is an accumulation of exposure to the news (Bogart, 1981). To the extent that coverage of "important" news within and across news channels is complementary or redundant, audiences are more likely to receive several accounts of the same news item—and that too might aid awareness and comprehension. Further, taking a week's worth of news virtually assures that a greater substantive variety of news will have time to unfold and develop. Increased story "maturity" may aid individuals seeking to integrate the news into semantic memory (Woodall et al., 1983). Finally, studying what people know about the news after a week may allow time for interpersonal discussion of the news. News stories sometimes serve as a "coin of social exchange" (Levy, 1978a, 1978b), and, as was suggested above, it is possible that the interpersonal utility of the news may enhance its recall and comprehension.

SAMPLES

Two separate probability surveys were conducted by telephone from the Survey Research Center at the University of Maryland using the random-digit-dialing method of household selection. In the first sample, 407 adults living in the greater Washington, DC metropolitan area were interviewed during May 1983. To maximize generalizability and to obtain a greater variety of news stories, a second set of interviews with 544 adults nationwide were carried out during June of 1983.

Most interviewing was done on weekends and was based on the stories that had received prominent media attention in the preceding seven days. Interviews were spread approximately equally throughout each month. Completion rates were 74 percent for the regional sample and 71 percent for the national sample, after 10 callbacks to establish whether the telephone numbers were actually of telephone households.[1] Data from both surveys were weighted by respondent sex, age, and education to ensure representativeness with 1980 Census Bureau figures for the region and the nation.

NEWS MEDIA USE QUESTIONS

In both surveys, respondents were asked the following media exposure questions: (1) "Over the past seven days, have you read any daily newspaper?" (2) "Over the past seven days, have you read any national news or commentary magazine?" (3) "Over the past seven days, have you listened to any news broadcasts on the radio?" (4) "Over the past seven days, have you watched any news programs on television?" Each "yes" response was coded "1"; each "no" was coded "0." Responses on these four items were summed for each respondent into an index of news exposure.

All respondents were also asked, "Did you talk with anyone about anything in the news last week?" If "yes," respondents were then asked, "About how many such conversations did you have?" and "Who did you talk with?" Additionally, respondents reporting conversations about the news were asked: "Can you remember any of the things you talked about?" and if so, "What was that?"

News stories discussed by respondents in the national sample were then coded by two trained research assistants and one of the authors to determine whether the item was "definitely," "maybe," or "not" one of the "most important" national or international stories of the preceding week. (A discussion of how the week's "most important" stories were identified follows below). Intercoder agreement on this measure was 87 percent.

While respondents were free to mention conversations about local, national, or international news items, our analysis focused on "cosmopolitan" news (Merton, 1949), that is, national and international stories of more than purely local interest. While knowledge of different types of news may have different correlates, and awareness and comprehension of local news may be even more influenced by word of mouth, at least one study that systematically examined *both* national and local news stories has found considerable similarity in the background and media use correlates associated with both types of news (Becker and Whitney, 1980).

---

[1] Estimates of public comprehension of the news presented below are, of course, subject to sample biases due to the elimination of nontelephone households and the nonresponse rate. Telephone interviewing may have excluded, for example, some poorer, older, and non-English-speaking respondents, while some people who are less interested in the news and less well informed about it may have disproportionately refused to be interviewed. Thus, although it should not affect our analysis of the *correlates* of news comprehension, it is likely that our point estimates of news comprehension may be marginally inflated, even though the data were weighted to match census figures for education, age, and sex distribution.

In addition, to assess and control for individual interest in the news, respondents in the national sample were asked, "How closely do you follow national and international news in the media? Would you say very closely, somewhat closely, or not very closely." Respondents were also asked standard demographic items (e.g., sex, age, education, etc.).

SELECTION OF NEWS COMPREHENSION ITEMS

All respondents were asked "awareness" and "comprehension" questions about the "most important" stories from the week preceding the interview. The selection of "most important" news stories represented a collective judgment by a rotating panel of five or more high-ranking news professionals who worked for *The Washington Post*, *USA Today*, Associated Press, CBS News, and ABC News. (For a more general discussion of this technique, see Robinson et al., 1982.) Each week this panel of journalists was asked: (1) What have been the "big stories" of the past week? (2) What were the one or two most important things ("main points") that the audience should have learned from those stories?

Examples of "important stories" and their "main points" include: an earthquake in California with no fatalities; the nationwide outbreak of AIDS and its effect on homosexuals, Haitians, and others; Congress limiting the "tax cut" out of concern over the federal deficit; U.S. Senator John Glenn announcing his presidential candidacy; the Pope visiting Poland to support Solidarity; the reelection of British Prime Minister Margaret Thatcher by a landslide; and the expulsion of the U.S. ambassador from Nicaragua for "spying" and the subsequent reaction of the Reagan administration.[2]

In choosing the final list of stories for inclusion in the week's survey, every effort was made to include only items that had received extensive coverage in both print and broadcast news media. Finally, the wording of survey questions was checked with the panel to insure appropriateness and accuracy. Audience awareness of any given news story was measured by a "yes–no" response to a question which took the generic form: "Did you hear or read anything last week about ——?" If "yes," the respondent was then asked a question, usually open-ended, which had been designed to test the degree to which the individual knew a given story's "main point" as defined by the panel of journalists.[3] Over the course of each four-week field period, we asked respondents in the regional sample about 23 different persons, events, or phenomena that had been in the news, and respondents in the national sample about 31 distinct persons, events, or phenomena.

---

[2] For a complete listing of news stories and their associated awareness and comprehension questions, see Robinson and Levy (1986).

SCORING OF NEWS COMPREHENSION ITEMS

Responses to the open-ended comprehension items were coded by the authors using a 9-point scale, developed in connection with earlier comprehension studies (Robinson and Sahin, 1984). Coding was by consensus, with initial intercoder agreement in excess of 90 percent. Scores on the comprehension scale ranged from 0 to 8. Responses which recalled either the "wrong" story or failed to give its "main point" were rated 0 or 1, while a failure to recall anything beyond the initial awareness of the story was coded 2. Vague recollections of the story were coded 3 or 4, depending on response specificity. The code of 4 was assigned if the respondent was able to recall correctly some detail about the story (e.g., the identity of a the principal person in the story), but not the story's "main point." A reply containing the correct "main point," however, was coded 5, with one point added to the basic score of 5 for any additional correct detail provided—up to a score of 8 if the respondent noted not only the "main point" but three additional, cogent story details.

Thus, for example, in stories about AIDS, respondents were asked, "Do you know about any particular group of people who have been affected (by AIDS) more the others?" A response of "homosexuals" was coded 5. If the respondent said "Homosexuals" and Haitians, drug addicts, and/or hemophiliacs, that response was scored as a 6, 7, or 8, depending on how many groups beyond the "main point" (homosexuals) were named.

A summary measure representing total news comprehension for the week was then constructed for each respondent.[4] For each answer to a specific comprehension item scored 5, 6, 7 or 8, the respondent received one point. These points were then summed and normalized to equalize the total number of weekly news items which varied slightly from week to week. Scores on this summary measure of news comprehension ranged from 0 to 9 in the regional sample and from 0 to 14 in the national sample. A score of 0 indicated no stories comprehended, while a summary score of 9 in the regional sample or 14 in the national study indicated comprehension of all items for a given week.

---

[3] Examples of "main point" comprehension questions include: "How many people were killed in the (California) earthquake?" "Why did Congress want to limit the tax cut?" "Why was Senator Glenn in the news?" "What do you think the purpose of the Pope's trip (to Poland) was?

[4] Obviously, two comprehension measures have been developed: the first gauges the "depth" of information on a single story; the second, the summary score, measures the

Ideally for our research, we should have had a specially devised "information meter" that would have automatically registered when respondents understood a news story and that recorded the source from which the information came. Given the current unavailability of such a device, we simulated the process with surveys. Our simulation asked first about the news information that respondents did have and second about the news sources respondents used. If users of news medium A possessed information that users of news media B did not, we took that to be reasonable evidence that news medium A was the more important information source. While this is not causal evidence (if indeed such social science data can be definitively gathered at all), it is the closest *field* approximation that we believed could be devised.

Results

The mean "awareness" score was 68 percent in the regional sample and 62 percent in the national sample. That means that roughly two-thirds of respondents in each sample said they had heard or read about the story. Details of U.S. government actions, in both domestic and foreign policy, tended to rank near the bottom in terms of public awareness, while respondents were most aware of especially dramatic news of real or potential danger and "human interest" stories.[5]

Means news comprehension scores were 2.7 (out of 9 weekly stories) in the regional sample, and 4.6 (out of an average of 14 stories per week) in the national sample. Thus, on average, somewhat less than one-third of the major news stories examined each week were comprehended by respondents sampled. No respondents in the regional sample successfully understood the "main point" of more than five of the week's major news stories. Fewer than 4 percent of national re-

---

"breadth" of respondent comprehension *across* a number of news stories. Limited analysis to date has found that the media use and interpersonal communication correlates of the two comprehension measures are highly similar.

[5] A story-by-story listing of "awareness" and "comprehension" scores is available on request from the authors.

Table 1. Weekly Conversations About News

	Regional Sample (407)	*National Sample (544)*
No. of conversations		
None	35%	40%
1–2	10	14
3–5	25	23
6–9	16	12
10 or more	14	11
Discuss "big" news		
Yes	n.a.	18
Maybe	n.a.	15
No	n.a.	51
Can't recall topic(s)	20	16

spondents comprehended 12 or more items and none had a score of 14. In general, these results are consistent with earlier parallel studies conducted in Great Britain (Robinson and Sahin, 1984); and like the pattern reported above for the awareness measure, comprehension was often, but not exclusively, higher for stories with dramatic or "human interest" content.

Almost two-thirds (65 percent) of respondents in the regional sample and almost as many (60 percent) of national respondents said they had talked "about things in the news" during the week prior to their being interviewed (Table 1). In both samples, the largest proportion of conversations, two-thirds or more, were with "friends" or "colleagues at work," with family members making up the remainder of conversational partners.

Only 20 percent of regional respondents and 16 percent of the national sample who reported having news-oriented conversations were unable to recall the topic of those discussions. This is a rather low proportion, based as it is on an unaided recall question. Indeed, it suggests that news—be it local, national, international, or some mix—has a relatively high salience for most individuals.

Of respondents (national sample only) who could recall what they talked about, some 21 percent reported having at least one conversation which was coded as "definitely" dealing with one of the "most important" stories of the week. Another 18 percent of the national sample reported at least one conversation that was coded as "maybe" focusing on one of the week's major stories.

BIVARIATE CORRELATES OF NEWS COMPREHENSION

The average scores on the summary measure of news comprehension for each major background variable group are presented in Table

Table 2. Mean News Comprehension Scores by Demographic, Media Use, and Communication Behavior Variables

	Regional Sample	*National Sample*
All respondents	2.7 (407)	4.6 (544)
Sex		
Male	3.3 (180)	5.7 (237)
Female	2.3 (226)	3.8 (307)
Age		
18–19	0.7 (9)	2.9 (17)
20–29	2.3 (97)	3.8 (127)
30–39	3.2 (90)	4.8 (113)
40–49	3.2 (63)	5.1 (82)
50–59	2.7 (68)	5.0 (63)
60–69	2.6 (49)	5.4 (76)
70 or older	2.7 (31)	4.2 (66)
Education		
< High school	1.7 (80)	3.1 (95)
H.S. grad	2.1 (119)	4.0 (278)
Some college	2.7 (71)	5.4 (94)
College grad	4.0 (137)	6.7 (49)
Graduate educ	—[a]	8.4 (28)
News exposure index		
0	0.5 (6)	1.2 (3)
1	1.0 (37)	2.8 (51)
2	2.1 (76)	3.6 (156)
3	2.8 (177)	5.0 (234)
4	3.6 (110)	6.2 (100)
How closely follow news		
Very closely	n.a.	6.4 (101)
Somewhat	n.a.	5.3 (266)
Not very	n.a.	2.5 (177)
Talks about news		
None	1.8 (143)	3.3 (219)
1–2	2.3 (38)	4.2 (74)
3–5	2.8 (101)	5.5 (127)
6–9	3.8 (66)	5.7 (64)
10 or more	3.9 (59)	6.8 (59)
Discuss "big" news		
Yes	n.a.	6.8 (98)
Maybe/no/DK	n.a.	4.1 (446)

[a] Respondents with post-college education included in college graduate category.

2. In general, there is considerable consistency between the regional and national samples with regard to these correlates of news comprehension. In both surveys, for example, there were fairly large differences in the comprehension scores of men and women, with men tending to have higher scores. Comprehension also increased with respondent age, although both sets of data showed some tendency for curvilinearity, with older respondents scoring lower than middle-aged respondents.

In both samples, level of formal education was the major predictor of news comprehension (see Robinson, 1967; Hyman et al., 1977 for similar results with other information survey questions). The average comprehension score of respondents with a college degree, for example, was more than double that of respondents who had graduated from high school.

Exposure to news media was also a clear but somewhat less powerful correlate of comprehension scores. For example, respondents in both samples with the least exposure to the news (i.e., 0 or 1 on the News Exposure Index) had mean scores that were less than half those achieved by those respondents reporting exposure to all four media news sources. However, because of the small size of this minimal exposure group, news exposure had less of a correlation with comprehension scores than did formal education. Interest in the news was another important correlate of comprehension scores, with respondents who claimed to follow the news "very closely" also scoring over twice as high as those who followed the news "not very closely."

Average news comprehension scores were also clearly correlated with the interpersonal communication variable. Talking about news with other people was related to considerably higher comprehension levels. As the number of conversations about news increased from none to 10 or more for the week, the average comprehension score more than doubled in both samples. Additionally, respondents in the national sample who reported talking about one or more major stories of the week had comprehension scores which were among the very highest in the study.

Table 2, then, links news comprehension scores with several possible predictor variables, considering each of those variables one at a time. The question naturally arises whether some of these correlates emerge as more powerful predictors after simultaneous adjustments are made for each of the other correlates in Table 2. For example, many of the correlates of comprehension examined here are also often linked with the strongest correlate of comprehension, namely, formal education. It could well be that controlling for the respondent's level of education would substantially reduce, or even eliminate, the correlations of news comprehension with news media exposure, interest in the news, or news-oriented conversations.

MULTIVARIATE ANALYSIS OF NEWS COMPREHENSION

News comprehension scores were next subjected to a Multiple Classification Analysis (Andrews et al., 1973) in order to determine the impact on comprehension of media use and communication behavior variables after controlling for demographic factors and the other media use questions. This MCA technique provides a combination multiple regression and analysis of variance adjustment on survey data.[6]

As Table 3 demonstrates, differences in the average comprehension scores for men and women in both samples are substantially reduced, after the other factors in Table 2 are taken into account. Statistical controls also reduce some of the curvilinearity of news comprehension and age in the regional sample, while educational differences and comprehension in both samples are also somewhat reduced.

Differences by levels of news media exposure are also considerably constricted by multivariate controls; and similarly, the gap is narrowed between respondents in the national sample who follow the news "very closely" and only "somewhat closely."

The interpersonal discussion variable is also affected by simultaneous controls for other variables, and several departures from the earlier pattern are found as well. Thus, in the regional sample, peak comprehension scores are reached in the 6–9 conversation group and then dip slightly in the 10 or more conversation category. In the national sample, peak comprehension scores are reached in the 10+ conversation group, while the mean comprehension score for respondents reporting 6–9 conversations dips below the score for the 3–5 conversation group.

The small sample sizes for the two groups reporting the largest number of conversations may account for some of the instability in the data. Indeed, if the data from the two samples in Table 3 were merged, the combined data would show clear monotonicity with number of conversations. The difference in average comprehension scores between the no-conversation group and the 10+ conversation group then would be 1.1 for the regional sample and 1.6 for the national sample, representing more than a 30 percent gain in news comprehension after MCA adjustments. For the national sample that is a larger difference than the 0.8 point gain between the four-media vs. 0–1 media use group combined (to obtain meaningful sample size); for the regional sample, the 1.4 point gain in mean comprehension scores across news media use is 0.3 points larger, but that is not a large difference.

---

[6] One advantage of MCA over other multiple regression techniques is that it provides straightforward adjusted averages for specific categorical groups in the sample (e.g., for men and women, or blacks, whites and other racial groups.) An MCA is also useful for ordered groups (e.g., high vs. middle vs. low education), since one can see whether relationships are linear and monotonic or have "peaks" and "valleys" in the middle categories. The principal drawback of MCA is that it does not provide error estimates that are necessary to conduct standard statistical tests on average difference scores between groups. However, certain "rules of thumb" have been developed for interpreting average group differences, and they are followed, where applicable, in this article.

Table 3. MCA Analysis of Comprehension Scores by Demographic, Media Use, and Communication Behavior Variables

	Regional Sample	*National Sample*
All respondents	2.7 (407)	4.6 (544)
Sex		
Male	3.0 (180)	5.3 (237)
Female	2.6 (226)	4.0 (307)
Age		
18–19	1.9 (9)	3.9 (17)
20–29	2.3 (97)	3.8 (127)
30–39	2.9 (90)	4.3 (113)
40–49	2.7 (63)	4.9 (82)
50–59	2.8 (68)	5.0 (63)
60–69	2.9 (49)	5.6 (76)
70 or older	3.7 (31)	4.8 (66)
Education		
< High school	2.0 (80)	3.4 (95)
H.S. grad	2.2 (119)	4.3 (278)
Some college	2.7 (71)	5.2 (94)
College grad	3.7 (137)	6.1 (49)
Graduate educ	—[a]	7.1 (28)
News exposure index		
0	2.0 (6)	3.3 (3)
1	1.6 (37)	4.1 (51)
2	2.4 (76)	4.3 (156)
3	2.9 (177)	4.8 (234)
4	3.1 (110)	4.8 (100)
How closely follow news		
Very closely	n.a.	5.6 (101)
Somewhat	n.a.	3.0 (266)
Not very	n.a.	3.5 (177)
Talks about news		
None	2.3 (143)	4.1 (219)
1–2	2.3 (38)	4.2 (74)
3–5	2.6 (101)	5.1 (127)
6–9	3.6 (66)	4.9 (64)
10 or more	3.4 (59)	5.7 (59)
Discuss "big" news		
Yes	n.a.	5.7 (98)
No	n.a.	4.4 (446)

[a] Respondents with post-college education included in college graduate category.

Thus, in both samples, the extent of discussion of the news seems to be at least as powerful a predictor of comprehension as the extent of news media exposure, and in the more generalizable national sample, such discussion was associated with almost twice as much spread in news comprehension as was media exposure.

A final insight into the importance of the interpersonal factor is given at the bottom of Table 3. If a respondent reported talking about even one of the week's "big news" stories, then, even after controlling for all other correlates reported in the table, that respondent was likely to score significantly higher on comprehension. Coupled with the findings reported above about comprehension and the number of news-oriented discussions, this result suggests that both the quantity *and* the quality of conversations are important factors in understanding the news.

Discussion

In summary, then, what can be said about the public's awareness and comprehension of national and international news over a week's time? First, and perhaps most obviously, it should be noted that there is a striking similarity between findings from the two surveys. Despite considerable geographic differences among respondents in the two samples and despite differences in the news stories examined each week, there is considerable overlap in the patterns within and the relative importance between key news comprehension correlates. This consistency across two independent samples points to a reassuring degree of generalizability in our findings.

Second, while education was found to be the single most important predictor of news comprehension, respondents who reported more conversations about the news were also more likely to score higher in news comprehension. Talking about the news, especially the week's "most important" national and international news, seems to have at least as powerful an effect on comprehension as mere exposure to the news media.

That finding, we believe, has serious implications for journalism, public understanding, and even the democratic process. While a full discussion of those implications is well beyond the scope of this article, we want to make one observation in passing. Journalists and politicians often complain that the public is not interested in the news and lacks the educational background to understand what gets reported (Robinson et al., 1982). Our study demonstrates, however, that controlling for news interest and education, public understanding of the news does increase as people talk about it. Journalists and others would do well, then, to consider ways to create and present news which stimulates or otherwise takes advantage of this powerful interpersonal "second stage" of the information flow.

Finally, the findings presented here bear directly on our general theories and models of the mass communication process. Two general paradigms inform most communication scholarship: one, the so-called dominant or American approach, is a fundamentally empirical and pluralistic model of relatively limited communication effects in which the power of the mass media is often mitigated by interpersonal influences; the other, the so-called critical or European perspective, finds major conceptual and empirical flaws in the dominant model and assumes often pervasive power for the media (Curran et al., 1982; Gitlin, 1978).

Our study, we believe, suggests a bridge between the two paradigms. On the one hand, our findings strongly suggest that interpersonal communication plays a crucial role in the reception, retransmission, and interpretation of mass media messages. That conclusion argues in support of the dominant theory and its continuing emphasis on the importance of social networks for the mass media process.

However, that same conclusion, namely that interpersonal communication increases the fidelity between the manifest content of mass media messages and the interpretations and meanings derived from those messages by audiences, can be interpreted to support the notion of a more powerful, "mainstreaming," or possibly even "ideological effect" of mass communication (Hall, 1983). From this perspective, interpersonal communication could be seen as helping to create audience "readings" of mass media messages that are closer in substance to "professional," and thereby ideologically "preferred," understandings of reality (Hall, 1980).

To substantiate such an assertion, of course, would require findings drawn from many different research sites in the mass communication process, and would necessarily involve both micro- and macro-studies of audiences, communicators, and culture more generally. This article obviously has a more limited scope, although our findings do have implications for such an investigation.

We have not, for example, offered *direct* evidence that the public interprets the news from any particular ideological standpoint; and we are aware that audiences are at least capable of creating "negotiated" or "oppositional" decodings of news messages (Morley, 1980). But we also have shown that public awareness and comprehension of the news, to the extent it occurs at all, often bears a striking similarity to journalistically encoded reality.

Similarly, the complex and often problematic relationship between journalistic encodings of the news and the more general phenomenon of ideology remains unexamined here. But other researchers have already shown that the ideological and thematic range of most news messages is fairly limited (Gans, 1979), and that represents an important precondition for demonstrating a linkage between audiences, communicators, and ideologies.

Taking these factors into account, and fully recognizing that much more work is needed, we still think it is plausible to speculate that conversations about the news will both increase the "accuracy" with which news messages are comprehended and the degree to which those messages may have ideological, essentially status quo consequences.

References

Andrews, F. M., J. N. Morgan, J. A. Sonquist, and L. Klem
1973 Multiple Classification Analysis. Second Edition. Ann Arbor: Institute for Social Research, The University of Michigan.
Becker, Lee, and D. C. Whitney
1980 "Effects of media dependencies." Communication Research 7:95–120.
Bogart, L.
1981 Press and Public. Hillsdale, NJ: Lawrence Erlbaum Associates.
Carter, R., and B. Greenberg
1965 "Newspapers or television: which do you believe?" Journalism Quarterly 42:29–34.
Chaffee, S.
1982 "Mass media and interpersonal channels: competitive, convergent, or complementary?" In G. Gumpert and R. Cathcart (eds.), Inter/media. Second Edition. New York: Oxford University Press.
Curran, J., M. Gurevitch, and J. Woollacott
1982 "The study of the media: theoretical approaches." In M. Gurevitch, T. Bennett, J. Curran, and J. Woollacot (eds.), Culture, Society and the Media. New York: Methuen.
Findahl, O., and B. Höijer
1976 Fragments of Reality: An Experiment with News and TV Visuals. Stockholm: Sveriges Radio.
Gans, H.
1979 Deciding What's News. New York: Pantheon.
Gantz, W.
1983 "The diffusion of news about the attempted Reagan assassination." Journal of Communication 33 (Winter):56–66.
Gerbner, G., L. Gross, M. Morgan, and N. Signorielli
1980 "The 'mainstreaming' of America: Violence Profile No. 11." Journal of Communication 30:10–29.
Gitlin, T.
1978 "Media sociology: the dominant paradigm." Theory and Society 6:205–253.
Greenberg, B.
1964 "Person to person communication in the diffusion of news events." Journalism Quarterly 41:489–494.
Gunter, B.
1980 "Remembering the television news: effects of visual format on information gain." Journal of Educational Television 6:8–11.
Hall, S.
1980 "Encoding and decoding." In S. Hall, D. Hobson, A. Lowe, and P. Willis (eds.), Culture, Media, Language. London: Hutchinson.
1983 "The rediscovery of 'ideology': return of the repressed in media studies." In M. Gurevitch, et al. (eds.), Culture, Society and the Media. New York: Methuen.
Hyman, H., C. Wright, and J. Reed
1977 The Enduring Effects of Education. Chicago: University of Chicago Press.
Katz, E., H. Adoni, and P. Parness
1977 "Remembering the news: what the picture adds to recall." Journalism Quarterly 54:231–239.
Katz, E., and P. Lazarsfeld
1955 Personal Influence. New York: The Free Press.
Levy, M.
1978a The Audience Experience with Television News. Association for Education in Journalism and Mass Communication, Journalism Monographs, Number 55.
1978b "Opinion leadership and television news uses." Public Opinion Quarterly 42:402–406.

Lichty, L.
1982 "Video vs. print." The Wilson Quarterly, Special Issue.
Mason, R.
1963 "The use of information sources by influentials in the adoption process." Public Opinion Quarterly 27:455–466.
Merton, R.
1949 "Patterns of influence: a study of interpersonal influence and communication in a local community." In P. Lazarsfeld and F. Stanton (eds.), Communication Research 1948–1949. New York: Harper and Brothers.
Morley, D.
1980 The 'Nationwide' Audience. London: British Film Institute.
Neuman, W. R.
1976 "Patterns of recall among television viewers." Public Opinion Quarterly 40:115–23.
Robinson, J.
1967 "World affairs information and mass media exposure." Journalism Quarterly 44:23–30.
1971 "The audience for national TV news programs." Public Opinion Quarterly 35:403–05.
1976 "Interpersonal influence in election campaigns: two-step flow hypotheses." Public Opinion Quarterly 40:304–19.
Robinson, J., and M. Levy
1986 The Main Source: Learning from Television News. Beverly Hills, CA: Sage Publications.
Robinson, J., and H. Sahin
1984 Audience Comprehension of Television News: Results from Some Exploratory Research. London: BBC Audience Research Department.
Robinson, J., H. Sahin, and D. Davis
1982 "Television journalists and their audiences." In J. Ettema and D.C. Whitney (eds.), Individuals in Mass Media Organizations. Beverly Hills: Sage Publications.
Rogers, E.
1983 Diffusion of Innovations. Third Edition. New York: Free Press.
The Roper Organization
1983 Trends in Attitudes Toward Television and Other Media: A Twenty-Four Year Review. New York: Television Information Office.
1985 Public Attitudes Toward Television and Other Media—A Time of Change. New York: Television Information Office.
Schramm, W.
1973 "Channels and audiences." In I. de Sola Pool and W. Schramm (eds.), Handbook of Communication. Chicago: Rand McNally.
Stevenson, R., and K. White
1980 "The cumulative audience of network television news." Journalism Quarterly 57:477–81.
Stone, V.
1969–1970 "Sources of most news: evidence and influence." Journal of Broadcasting 14:1–4.
Tichenor, P., J. Rodenkirchen, C. Olien, and G. Donohue
1973 "Community issues, conflict, and public affairs knowledge." In P. Clarke (ed.), New Models for Communication Research. Beverly Hills, CA: Sage Publications.
Troldahl, V., and R. Van Dam
1965 "Face to face communication about major topics in the news." Public Opinion Quarterly 29:626–34.
Woodall, W., D. Davis, and H. Sahin
1983 "From the boob tube to the black box: television news comprehension from an information processing perspective." Journal of Broadcasting 27 (Winter):1–23.

Measuring Change in Personal Economic Well-Being

STEVEN J. ROSENSTONE, JOHN MARK HANSEN
AND DONALD R. KINDER

FOR OVER A decade, researchers have tried to unravel the political consequences of economic conditions—with mixed success. Studies using aggregate, over-time, data have found that prosperity enhances the electoral chances of incumbents and bolsters support of chief executives.[1] Parallel investigations at the individual level, however, have turned up a striking anomaly: changes in citizens' personal economic well-being seem to have little or no bearing on their political preferences.[2]

Efforts to reconcile these findings have ventured down several paths. Perhaps citizens respond not to changes in their own personal economic situation but to evaluations of the economic health of the nation as a whole (Kinder and Kiewiet, 1979, 1981) or the economic

---

Abstract Individual-level investigations of the link between personal economic well-being and political preference commonly turn up findings at odds with aggregate, over-time analyses. Although various theoretical explanations for this anomaly have been advanced, we focus on a different culprit: the possibly frail measure employed in survey based studies. We develop improved measures of change in personal economic well-being and use a national survey of the American electorate to test them. Our findings are striking: although the traditional question used by survey researchers is quite valid, its unreliability has caused the effect of personal economic well-being on political evaluations to be substantially underestimated.

Steven J. Rosenstone is Professor of Political Science at Yale University. John Mark Hansen is Assistant Professor of Political Science at The University of Chicago and is a Research Fellow at The Brookings Institution. Donald R. Kinder is Associate Professor in the Department of Political Science and Associate Research Scientist at the Institute for Social Research, both at the University of Michigan.

[1] E.g., Kramer, 1971; Bloom and Price, 1975; Frey and Schneider, 1975; Tufte, 1978; Kernell, 1978; Hibbs and Fassbinder, 1981; Hibbs, 1982; and Rosenstone, 1983.

[2] E.g., Fiorina, 1978, 1981; Kinder and Kiewiet, 1979, 1981; Schlozman and Verba, 1979; Sears et al., 1980; Kinder, 1981; Kiewiet, 1983.

Public Opinion Quarterly Vol. 50:176–192 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-176/$2.50

predicaments of groups (Conover, 1985; Kinder et al., 1983). Perhaps personal finances have political consequences, but only for rather unusual voters (Feldman, 1982; Kinder and Mebane, 1983). Or perhaps the political effect of family finances has to do not with the voter's calculus but with the caliber of candidates: the appearance of attractive, well-financed out-party candidates hoping to reap windfalls from the incumbents' economic woes (Fiorina, 1978; Jacobson and Kernell, 1981).[3]

We take a different route. We argue here that the frail results produced by survey-based analyses reflect in part the frail *measurement* of changing personal economic well-being. While relatively trustworthy measures of aggregate economic conditions are readily available, the reliability and validity of individual-level measures of personal financial security are largely untested. Most survey work, in fact, has depended on answers to a single question asked (with slight variation) by the National Elections Studies since 1952:

> We are interested in how people are getting along financially these days. Would you say that you (and your family living with you) are better off or worse off than you were a year ago?

Pinpointing the connection between economic circumstance and political preference based on this lone item is a perilous enterprise.[4]

Motivated by the theoretical issues at stake, we report here our efforts to improve the measurement of change in personal economic well-being. We attend to two kinds of measurement error: those arising from use of a single item rather than a set, and those deriving from use of a time frame (one year) that may be beyond respondents' memories. We draw upon a national survey conducted in the summer of 1983 to assess the adequacy of the traditional better/worse off question and compare its performance to new items we designed. While our results have immediate implications for the proper measurement of personal economic well-being, they also hold lessons for the measurement of subjective states more generally and for our understanding of the interplay between economics and politics in the lives of ordinary citizens.

---

[3] Kramer (1983), in addition, has argued that individual-level findings are statistical artifacts and that the mechanisms linking economic conditions to politics can never be uncovered using cross-sectional data.

[4] The better/worse off financially question is sometimes supplemented by others: by measures of unemployment and underemployment (e.g., Fiorina, 1981; Kinder and Kiewiet, 1979), by questions that invite citizens to name their most pressing personal problems, which are often economic (e.g., Kinder and Kiewiet, 1981; Kiewiet, 1983), and in at least two instances by sets of questions that attempt to map personal economic experiences and judgments more comprehensively (Kinder and Mebane, 1983; Mebane, 1982).

Measuring Change in Personal Economic Well-Being

THE CONCEPT

Because we cannot observe change in personal economic well-being directly, we asked respondents to report on perceptions and activities that change as a consequence of fluctuations in their own well-being. We assume each perception or activity—which is observable—is an indicator of change in economic well-being. We assume further that a person may cope with economic adversity in any of five ways. She can consume less or buy less expensive goods, look for more work, borrow, save less or dip into savings, or put off paying creditors. As her economic well-being improves, on the other hand, she can consume more, work less, forgo borrowing, save more, and pay debts promptly. Our survey questions tap each of these consequences of change in well-being. We also ask for a global assessment of whether family income has increased or decreased. Finally, we expand the standard better/worse off financially question by adding a follow-up that asks people how much better or worse off they have become.[5]

TIME FRAME

The traditional better/worse off financially question asks people to compare their current situation to last year's. Two considerations motivate us to consider a shorter time frame in this item as well as in our new ones. First, responses to questions asking for six-month instead of one-year comparisons may be more reliable. People rapidly forget even tangible and salient events like hospitalizations, automobile accidents, crime victimizations, and major home improvement expenditures (Cannell et al. (nd.); Penick, 1976; Neter and Waksberg, 1964). About *half* such instances are not reported to survey interviewers a year after their occurrences. Second, short-term changes in personal economic well-being may explain political preferences and behavior better than changes over a longer term. In evaluating presidents or deciding how to vote, people may ask, "What have you done for me lately?" and really mean lately.[6]

---

[5] Our conception of personal economic well-being does not include specific reference to unemployment. Unemployment is conceptually distinct from change in personal economic well-being. The two are related, to be sure, but a family can experience unemployment without experiencing economic adversity and can experience economic adversity without experiencing unemployment.

[6] Fair (1978) found that a six-month change in real GNP per capita is a slightly better predictor of presidential voting than a one-year change; Rosenstone (1982) found that short-term changes in unemployment suppress turnout more than long-term changes; Kernell (1978) found that six-month moving averages for differences in unemployment and consumer prices best explain fluctuations in presidential popularity.

ALTERNATIVE MEASURES

The survey questions we devised were included in the Center for Political Studies Pilot Survey for the 1984 National Election Study. The 314 people interviewed in July 1983 and reinterviewed a month later were a subsample of respondents to the 1982 National Election Study, a random sample of citizens of voting age.[7] To test the relative merits of the two time frames, respondents were randomly assigned to one of two groups. Each group was asked the same personal well-being questions, with one difference. People in sample A ($N = 158$) were prompted with phrases like "in the last year" or "this past year" while those in sample B ($N = 156$) were prompted with "in the last six months" or "this past six months."

The questions and the marginals for the two samples are reported in Table 1.[8] The very few people (.1 to 2.0 percent) who responded "don't know" or whose answers were not ascertained are deleted from the percentages displayed in the table. In general, the six-month time frame elicited slightly "better off" responses than the one-year prompt. This is to be expected: compared to the previous year, the economy really was better in the six months prior to the interviews. The rate of growth in per capita real disposable income between January 1983 and July 1983 was five times greater than between July 1982 and July 1983 (245 vs. 55 percent). Unemployment dropped .9 percent in the six months prior to the interviews, but fell only .3 percent over the previous year. Thus, it seems respondents did indeed match their recall to our time frames. We discuss other evidence on the discriminant validity of the two frames later.

MEASUREMENT MODEL

We employ the familiar Jöreskog (1970) maximum likelihood factor analysis model available in LISREL V to uncover the dimensions—perhaps more than one—that underlie our various measures of change in economic well-being. (For convenience, we coded all variables to a zero-one interval, with 1.0 representing the "better-off" responses.) Because we need to compare parameter estimates for the one-year and six-month time frames, we analyzed the two samples separately, estimating the coefficients from the variance-covariance matrix among the indicators.

---

[7] These data, as well as a description of sampling methods, are available from the ICPSR.

[8] The variable numbers reflect the order in which the questions were asked. Items 1–4 appeared as the opening questions on the questionnaire; the rest were asked three minutes later. Between the two sets of items respondents were asked about the national economy.

Table 1. Survey Questions Asked to Measure Changes in Personal Economic Well-Being—Variable Number, Question, and Marginals (in Percents)

Variable Number	*Question*	*Sample A 1 Year Time Frame*	*Sample B 6 Month Time Frame*
V1	We are interested in how people are getting along financially these days. Would you say that you (and your family living with you) are better off or worse off than you were (a year ago/six months ago)?		
	Better off	30.4	32.1
	Same	32.9	39.7
	Worse off	36.7	28.2
V2	Is that much better off or somewhat better off? Is that much worse off or somewhat worse off?		
	Much better off	7.6	7.1
	Somewhat better off	22.8	25.0
	Same	32.9	39.7
	Somewhat worse off	26.6	21.2
	Worse off	10.1	7.1
V3	During the (past year/past six months) have you (and your family) been able to buy most of the things you needed and planned on, or have you had to put off buying these things?		
	Had to put off buying	53.2	51.6
V4	Do you think that over the (last year/past six months) (your/your family's) income has gone up more than the cost of living, has it fallen behind, or has it stayed about even with the cost of living?		
	Gone up more	10.1	8.5
	Stayed about even	45.6	47.4
	Fallen behind	44.3	43.8
	Now I'm going to read a list of things that may have happened to you (or family members living with you) during the past (year/six months).		
V5	(This past year/In the last six months) have you (or anyone in your family) had to watch your budget more closely than the (year before/six months before)?		
	Yes	63.9	55.1
V6	During the past (year/six months), have you (and your family) put off medical or dental treatment because you didn't have the money?		
	Yes	29.7	27.6
V7	(This past year/In the past six months), in order to make ends meet, did (any of) you borrow money from a bank, lending institution, or from relatives or friends?		
	Yes	25.9	18.1

Table 1. *(Continued)*

Variable Number	*Question*	*Sample A 1 Year Time Frame*	*Sample B 6 Month Time Frame*
V8	[If yes] Did you have to borrow more, the same amount, or less than the (year/six months) before or didn't you have to borrow then?		
	More	3.8	6.5
	Same	5.7	3.9
	Less	5.7	1.3
	Didn't borrow then	10.8	6.4
V9	[Change in borrowing] [Constructed from V7 and V8]		
	Not borrowing	74.1	81.9
	Borrowing less	5.7	1.3
	Borrowing same	5.7	3.9
	Borrowing more	14.6	12.9
V10	(This past year/In the past six months) have you (and your family) had to use your savings to make ends meet?		
	Yes	47.5	37.8
V11	[If yes] Did you use your savings more, the same, or less than the (year before/six months before), or didn't you have to use your savings then?		
	More	30.4	12.3
	Same	5.7	13.5
	Less	1.9	2.6
	Didn't use savings then	9.5	9.0
V12	[If no] Over the past (year/six months) have you (and your family) been able to put money aside?		
	Yes	31.2	37.2
V13	[If yes] Have you been able to save more, the same amount or less than (the year before/during the six months before)?		
	More	10.2	10.3
	Same	15.9	20.5
	Less	5.1	6.4
V14	[Change in savings] [Constructed from V10, V11, V12, and V13]		
	Less savings	45.2	27.7
	Same savings	42.7	59.4
	Greater savings	12.1	12.9
V15	(This past year/In the past six months), in order to make ends meet, have (any of) you looked for a new job, or looked for a second job, or tried to work more hours at your present job?		
	Yes	44.9	34.2

As the estimates reported in Table 2 show, two distinct but correlated dimensions underlie the pool of indicators.[9] The first dimension produces economic evaluations: answers to the better/worse off financially, change in income, watch budget more closely, and change in savings questions. The remaining items—putting off buying things, putting off medical treatment, changing borrowing habits, and looking for a new job or more hours—are all reports of activities performed by people adjusting to economic changes. The evaluative dimension is comprised of things that people *think* about as a result of their financial states; the behavioral dimension is comprised of things that people *do* as a consequence of their financial states. (For similar results see Kinder and Mebane, 1983.) This model fits the data well regardless of whether the six-month or one-year time frame is employed. The adjusted measure of goodness of fit is .937 in both samples; the Chi-square has a probability of .657 in sample A and .638 in sample B.[10]

There is surely nothing in the measurement model that allows us to choose between the six-month and one-year time frames. The structure of the measurement model is precisely the same across the two samples; the loadings are nearly the same; the individual items are of comparable reliability; the two latent variables are correlated to the same degree; the estimated reliabilities of the indicators are virtually identical. We must search elsewhere for a basis to choose between the two time frames. But looked at in a more positive way, the results displayed in Table 2 provide striking confirmation of the robustness of the measurement model: the two dimensions of change in personal economic well-being show up in exactly the same manner in each of the two independent samples.

Construct Validity

We now have in hand four ways to measure personal economic well-being: the traditional three-category better/worse off financially ques-

---

[9] A blank cell indicates that the coefficient is statistically equivalent to zero. We initially treated each survey question as an indicator of a single latent dimension, but this specification fits the data poorly: prob. = .042 in sample A; prob. = .031 in sample B.

[10] We tested some alternative specifications. One allowed the savings question (V14) to load on both the evaluative and behavioral dimensions, but the item is apparently more an evaluation than a behavior. Saving does not really require that a person *do* anything; it is simply what is left in the checkbook at the end of the month. Borrowing, by contrast, involves a conference with a bank loan officer or a rich uncle. There was also no evidence of a response set within each of the two batteries of questions (V2–V4 and V5–V15). Finally, because the eight indicators are discrete (each has two to five response categories), we estimated the parameters using polychoric correlations, but found that the structure of the model and the relative loadings of the variables remained about the same. For details, see Rosenstone et al., (1983).

Table 2. Maximum Likelihood Factor Analysis of Items Measuring Changes in Personal Economic Well-Being (Estimates Made from Variance-Covariance Matrix)

		Loadings				
		Evaluative Dimension		*Behavioral Dimension*		
Variable	*Question*	*Coef*	*SE*	*Coef*	*SE*	*Reliability*
	Sample A—1 Year Time Frame					
V2	Better/worse off financially	.193	(.021)			.495
V3	Put off buying things			.334	(.041)	.446
V4	Change in income	.251	(.024)			.588
V5	Watch budget more closely	.337	(.037)			.490
V6	Put off medical treatment			.310	(.038)	.457
V9	Change in borrowing			.163	(.032)	.193
V14	Change in savings	.214	(.027)			.393
V15	Look for new job/more hours			.264	(.042)	.281
						Total .898
Correlation between the two dimensions = .781						
Chi-square with 19 d freedom = 16.01 (prob = .657)						
Adjusted goodness of fit = .937						
	Sample B—6 Month Time Frame					
V2	Better/worse off financially	.184	(.020)			.527
V3	Put off buying things			.378	(.042)	.567
V4	Change in income	.214	(.025)			.457
V5	Watch budget more closely	.366	(.038)			.538
V6	Put off medical treatment			.273	(.038)	.370
V9	Change in borrowing			.164	(.031)	.217
V14	Change in savings	.172	(.026)			.305
V15	Look for new job/more hours			.193	(.042)	.165
						Total .894
Correlation between the two dimensions = .759						
Chi-square with 19 d freedom = 16.29 (prob = .638)						
Adjusted goodness of fit = .937						

tion, the five-category version of the same question, the evaluative dimension, and the behavioral dimension.[11] We analyze two aspects of the convergent validity of each. First, how well is each predicted by

---

[11] We constructed a scale for the evaluative dimension and one for the behavioral dimension. The individual items were weighted by the factor scores generated from analyzing the variance-covariance matrix among the items comprising that dimension. (This was done separately for each dimension within each sample.) The scales are highly reliable:

	Sample A (1-Year)	Sample B (6-Month)
Reliability of evaluative scale	.81	.78
Reliability of behavioral scale	.70	.69

variables, like objective economic circumstances (changes in income, job loss) and social location (education, race, sex), that *should* be associated with changes in economic security? Second, do these associations vary in sensible ways with the time frames? To find out, we regressed each of the four measures of personal economic well-being on these validating variables and examined the unstandardized bivariate regression coefficients.

We found strong convergent validation for all four measures. Consider just a few examples. (For details see Rosenstone et al., 1983.) First, regardless of the measure of change in financial well-being used or the time frame employed, men are better off than women, whites better off than blacks, and the well educated are better off than those with less formal schooling. Second, unemployment in July 1983 is more highly associated with change in personal economic well-being measured over the last six months than over the last year. Third, change in family income between November 1982 and July 1983 is more strongly associated with one-year measures of change in personal economic well-being than with six-month measures.

One striking feature of these results is how well the traditional better/worse off financially item stands up to its competitors. The relationships between the various measures of objective economic circumstance and social location, on the one hand, and personal economic well-being, on the other, are about as strong when the latter is measured by the usual three-category format as when it is measured either by the five-category rendition or by the four items comprising the evaluative dimension scale. The traditional question is a respectably valid measure of change in personal economic well-being.

This point is reinforced by one final bit of validating evidence we gleaned from monthly data gathered by the Survey of Consumer Attitudes. We regressed responses to the traditional three-category better/worse off financially question with a one-year time frame onto monthly changes in aggregate per capita real disposable income. We estimated a series of equations, varying the number of months over which we calculated the change in real income. If people really do think back a full year as prompted, responses should be most highly associated with 12-month changes in income. As it happens, replies to the better/worse off financially questions are best predicted by changes in per capita real disposable income that are slightly longer than one year. The fit is best, and the coefficient is largest at about 14 months. (For details see Rosenstone et al., 1983.) Thus, people do seem to employ a baseline that is pretty close to the one they are asked to use.

In summary, the various measures of change in personal economic well-being, including the traditional three-category better/worse off financially question, perform well in our first set of validation tests. Each is predicted in expected ways by measures of objective economic circumstance and location in the social order. Moreover, various results suggest that respondents who are prompted to think back over the previous six months in fact do so; those prompted with a one-year frame look back over something close to one year. Personal economic well-being can indeed be measured adequately—a most important point—but these results do not single out the best measure among the several competitors. For that we must investigate the political consequences of personal economic well-being, our next and final test.

Predictive Validity

Here we examine, across the two time frames, how well the alternative measures predict three dependent variables: evaluations of the nation's economy, evaluations of Ronald Reagan's performance as president, and vote choices in a Reagan–Mondale trial heat (Tables 3–5).[12] For each dependent variable, we estimate a series of equations. The first equation employs the three-category better/worse off financially question as the measure of change in economic well-being; the second equation uses the five-category version; the third, the evaluative dimension; and the fourth, both the evaluative and the behavioral dimensions. The control variables are listed in a note at the bottom of each table.[13] We treat the financial well-being variables as endogenous and use two-stage least squares (2SLS) to correct for measurement error.[14]

We estimate a separate coefficient for the effect of change in personal economic well-being in each sample. By comparing the one-year

---

[12] We also estimated the effect of economic well-being on thermometer evaluations of Senator John Glenn to test the discriminant validity of the four measures. We expected the coefficients to be zero. They were.

[13] These demographic and political variables can safely be regarded as exogenous. We employed November 1982 measures of party identification, liberalism-conservatism, union membership, and family income to eliminate concerns about mediating effects of retrospective evaluations (Fiorina, 1981). The control variables are constrained to have the same effect in both samples.

[14] The causes of each measure of personal economic well-being discussed in the last section and the responses to the better/worse off financially question given in the November 1982 survey are used as exogenous instruments. The reduced form equations have sufficiently high R^2s—.29 to .44 for endogenous variables constrained to be equal across equations, .59 to .82 for variables unconstrained across equations—to yield reliable second-stage estimates.

Table 3. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Evaluations of the Nation's Economy—Two-Stage Least Squares Estimates[a]

		Coefficients			
Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Sample A (1 Year)*	*Sample B (6 Months)*	*Prob. Coef A = Coef B*	*Standard Error of Regression*
1.	Better/worse off financially—3 categories	.135 (.076)	.214 (.074)	.03	.224
2.	Better/worse off financially—5 categories	.228 (.110)	.314 (.106)	.06	.221
3.	Evaluative dimension	.294 (.113)	.337 (.105)	.25	.221
4.	Evaluative dimension	.080 (.200)	.355 (.279)		.224
	Behavioral dimension	.240 (.145)	.099 (.184)		
5.[b]	Evaluative dimension	.280 (.180)			.222
	Behavioral dimension	.136 (.118)			
6.[c]	Evaluative dimension	.433 (.218)	.549 (.213)		.236 (A) .212 (B)
	Behavioral dimension	−.038 (.191)	−.066 (.151)		

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; log (family income in November); and Hispanic. These variables are assumed to be exogenous. (In addition, dummy variables for rural residents, women, and unemployed respondents or mates appear as exogenous variables in equations 1 and 2, but were deleted [prob. < .65] from equations 3–6. Whether the respondent or his mate was unemployed during the year appeared as a variable in equations 1–4, but was deleted from equations 5 and 6 for the same reason.) The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

[c] These coefficients were estimated separately on each sample.

and the six-month coefficients, we can assess the predictive power of each question format. To help us in this task, we report the probability that the two coefficients are equal.[15]

---

[15] The coefficients in the fourth equation for each dependent variable—that which pits the behavioral dimension against the evaluative dimension—are too unstable to allow a meaningful comparison of the relative effects of the two dimensions. We are forced, therefore, to assume that the effects are constant across the two samples; these estimates are reported in equation 5. Finally, because evaluations of the nation's economy also were asked with both a one-year and a six-month time frame, it may be inappropriate to assume that the effects of the other exogenous variables are equal in the two samples. For this variable, then, we estimate separate equations for each sample (reported in equation 6, Table 3 only).

Table 4. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Evaluations of Reagan's Handling of His Job as President—Two-Stage Least Squares Estimates[a]

Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Coefficients*		*Prob. Coef A = Coef B*	*Standard Error of Regression*
		Sample A (1 Year)	*Sample B (6 Months)*		
1.	Better/worse off financially—3 categories	.436 (.114)	.323 (.104)	.22	.303
2.	Better/worse off financially—5 categories	.584 (.155)	.479 (.144)	.09	.303
3.	Evaluative dimension	.698 (.166)	.525 (.139)	.03	.300
4.	Evaluative dimension	.489 (.310)	.639 (.298)		.301
	Behavioral dimension	.116 (.206)	−.079 (.207)		
5.[b]	Evaluative dimension	.607 (.238)			.303
	Behavioral dimension	.007 (.161)			

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; union household in November; log (family income in November); education in November; race; Jew; and age. These variables are assumed to be exogenous. (Whether the respondent or his mate was unemployed in the last year also appeared as an exogenous variable in equations 1 and 2, but was deleted from equations 3–5 because its probability of being different from zero fell to .65.) The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

THREE VS. FIVE CATEGORY BETTER/WORSE OFF FINANCIALLY QUESTION

Our results are unequivocal and entirely startling. Simply by expanding the traditional measure's three response categories to five, we find sizable increases in estimated political effects. The increases show up in *all* comparisons: for evaluations of the national economy, ratings of President Reagan's performance, and vote choices in a Reagan–Mondale trial beat, and for six-month and one-year time frames alike. Put the other way around, in these data, *employing the traditional better/worse off financially question causes the political effect of changing personal economic conditions to be underestimated by as much as one-half.* For example, the estimated impact of changes in personal economic well-being on evaluations of Reagan's performance as president, after correcting for measurement error, is .436 using the

Table 5. Estimated Effect of Alternative Measures of Changes in Personal Economic Well-Being on Reagan-Mondale Trial Heat—Two-Stage Least Squares Estimates[a]

Eq #	*Measure of Change in Personal Economic Well-Being Appearing in the Equation*	*Coefficients* *Sample A (1 Year)*	*Sample B (6 Months)*	*Prob. Coef A = Coef B*	*Standard Error of Regression*
1.	Better/worse off financially—3 categories	.462 (.157)	.424 (.152)	.27	.397
2.	Better/worse off financially—5 categories	.627 (.220)	.612 (.215)	.44	.392
3.	Evaluative dimension	.463 (.230)	.421 (.205)	.37	.386
4.	Evaluative dimension	.824 (.446)	.317 (.430)		.394
	Behavioral dimension	−.226 (.297)	.059 (.287)		
5.[b]	Evaluative dimension	.503 (.330)			.389
	Behavioral dimension	−.055 (.214)			

[a] The other variables that appeared in each equation were: party identification in November; liberalism-conservatism in November; union household in November; log (family income in November); education in November; and Hispanic. These variables are assumed to be exogenous. The measures of personal economic well-being are treated as endogenous.

[b] In this equation, the coefficients were constrained to be equal across the two samples.

three-category measure but .584 when the five-category version is adopted (Table 4).[16]

EVALUATIVE DIMENSION VS. THE BETTER/WORSE OFF FINANCIALLY QUESTIONS

The evaluative dimension sometimes outperforms the five-category version of the better/worse off financially question and sometimes does not. It always outperforms the traditional three-category format. Thus, the marginal gain from employing the four-item scale over the five-category better/worse off financially item is not nearly as great as the gain from using the five- instead of the three-category better/worse off financially question.

---

[16] Had we failed to correct for measurement error, the coefficients, particularly for the traditional measure, would of course be attenuated still further. If, for example, we had used ordinary least squares instead of 2SLS, the coefficients for the one-year time frame in the first three equations listed in Table 4 would have been .271, .365, and .455 instead of .436, .584, .698.

THE EVALUATIVE VS. THE BEHAVIORAL DIMENSION

The behavioral dimension has no independent explanatory power over and above that of the evaluative dimension (look at equations 4 and 5 in each table).[17] This finding is not an artifact of the high association between these two dimensions; relatively efficient estimates are produced once the effects in the two time frames are constrained to be equal (in equation 5).[18]

THE SIX-MONTH VS. THE ONE-YEAR TIME FRAME

The contest between the two time frames does not produce a decisive winner. The six-month time frame is more strongly associated with evaluations of the nation's economy; the one-year time frame is more strongly related to evaluations of Reagan's performance as president, and choices between Reagan and Mondale. Few of these differences are sure bets: most of the probabilities fall short of conventional standards of statistical significance.

Conclusions

Where others have proposed theoretical solutions to the puzzling divergence between aggregate and individual-level findings on the political consequences of economic fluctuations, we have proposed a methodological one. The single survey item traditionally used to prove (or disprove) the link between changing economic well-being and political evaluations and behaviors, although valid, is quite unreliable relative to an expanded version of the same question and relative to the evaluative scale built from four items. The empirical consequences in our data are striking: with new measures, changes in respondents' personal economic well-being exercise substantial influence over their assessment of the national economic scene, the support they give their president, and vote choice.[19] Reliance on the traditional item alone seriously underestimates the political impact of changing personal economic well-being. In some degree, then, our analysis illustrates the perilous consequences that follow from the violation of an altogether familiar maxim. *Items that produce greater variance in responses are preferred to items that produce less variance.* Surely, our most startling result is the sharp boost in predictive validity that occurred when the five-category version of the better/worse off question replaced the traditional three-category version.

---

[17] This conclusion also holds when the five-category better/worse off financially question is used in lieu of the evaluative dimension in equation 5 in Tables 4 and 5, and in equation 6 in Table 3.

[18] We suspect that the behavioral dimension is a cause of the evaluative dimension. The behavioral scale is more strongly associated with unemployment than either the worse off financially questions or the evaluative scale. Moreover, variables measuring experience with unemployment drop out when the behavioral dimension is added to the equations for evaluations of the nation's economy and evaluations of Reagan's job performance.

[19] Sears and Lau (1983) argue that the political effect of change in personal economic well-being stems from asking respondents about their economic circumstances either immediately before or after soliciting their political judgments. Interviewers either inadvertently prime respondents with personal economic reasons for their political judgments, or respondents rationalize their political judgments in personal economic

Another striking result is the relatively poor performance of the behavioral dimension. Perhaps this is because the opportunities to engage in the activities that comprise the behavioral dimension vary across the population. Not all people cope with economic fluctuations in the same way. The elderly, we discovered, are less likely to put off buying things than to postpone medical and dental care. Blacks, perhaps reacting to realistic assessments of the job market, are less likely to look for work than to borrow, forgo medical treatment, or defer purchases. The variables that make up the evaluative dimension, on the other hand, are not subject to selective opportunities. Restriction of variability to certain subgroups may account for the behavioral dimension's lower reliability and weaker explanatory power.

It may also be that questions that ask for recall of specific instances are generally less reliable than questions that ask for comparisons of current states to past baselines. An anecdote will make the point. Nineteen respondents had said in their November 1982 interviews that either they or their spouses were unemployed at that time. Eight months later, in July 1983, only 8 of the 19 (42 percent) said that they or their spouses were unemployed *in the past year*. Assuming that there is little overreporting of current unemployment, this is a stunning demonstration of the rapidity with which memories of specific facts or behaviors lapse, a finding well in keeping with other studies of recall cited earlier. Although people find it difficult to recall isolated events (as in the items comprising the behavioral dimension), our analysis suggests that people can capably draw comparisons with recalled baselines (as in the questions that make up the evaluative dimension).

Although our results do not settle the theoretical debates that served as our point of departure, they do send some important warnings to the

---

terms. Our results stand up well to their arguments. First, 2SLS purges the coefficients of any effect of rationalization. Second, although questions about the national economy and Reagan's performance were asked in close proximity to the personal economic well-being questions, the Reagan–Mondale trial heat question appeared in an altogether different interview: it led off the second wave, some four weeks later. Change in personal economic well-being is no worse a predictor of trial heat choice than of the other two variables.

protagonists. Surely the debates cannot be decided conclusively in the absence of better data. The construct "change in personal economic well-being" is but one character in the theoretical cast; better measures of "change in national economic well-being" and "change in group economic well-being," to name two, must be developed as well. Nor is adequate measurement the whole battle. One customary and likely consequential difference between aggregate and survey studies of political economy is that the economic context varies in the first but not the second (Kramer, 1983). When over-time variance is introduced into survey-based studies along with adequate measures of personal, group, and national economic well-being, then the individual-level mechanisms that link economic changes to political outcomes may be more convincingly demonstrated.

References

Bloom, H. S., and H. D. Price

1975 "Voter response to short-run economic conditions: the asymmetric effect of prosperity and recession." American Political Science Review 69:1240–54.

Cannell, C. F., K. H. Marquis, and A. Laurent

n.d. "A summary of studies of interviewing methodology." Unpublished manuscript, Survey Research Center, University of Michigan.

Conover, P. J.

1985 "The impact of group economic interests on political evaluations." American Politics Quarterly 13:139–66.

Fair, R. C.

1978 "The effect of economic events on votes for president." Review of Economics and Statistics 60:159–73.

Feldman, S.

1982 "Economic self-interest and political behavior." American Journal of Political Science 26:446–66.

Fiorina, M. P.

1978 "Economic retrospective voting in American national elections: a microanalysis." American Journal of Political Science 22:426–43.

1981 Retrospective Voting in American National Elections. New Haven: Yale University Press.

Frey, B. S., and F. Schneider

1975 "On the modeling of politico-economic interdependence." European Journal of Political Research 3:339–60.

Hibbs, D. A., Jr.

1982 "The dynamics of political support for American presidents among occupational and partisan groups." American Journal of Political Science 26: 313–32.

Hibbs, D. A., Jr., and H. Fassbinder

1981 Contemporary Political Economy. Amsterdam: North-Holland.

Jacobson, G. C., and S. Kernell

1981 Strategy and Choice in Congressional Elections. New Haven: Yale University Press.

Jöreskog, K. G.

1970 "A general method for analysis of covariance structures." Biometrika 57:239–51.

Kernell, S.
1978 "Explaining presidential popularity." American Political Science Review 72:506–22.
Kiewiet, D. R.
1983 Macroeconomics and Micropolitics: The Electoral Effects of Economic Issues. Chicago: University of Chicago Press.
Kinder, D. R.
1981 "Presidents, prosperity, and public opinion." Public Opinion Quarterly 45:1–21.
Kinder, D. R., and D. R. Kiewiet
1979 "Economic discontent and political behavior: the role of personal grievances and collective economic judgments in congressional voting." American Journal of Political Science 23:495–527.
1981 "Sociotropic politics: the American case." British Journal of Politics 11:129–61.
Kinder, D. R., and W. R. Mebane, Jr.
1983 "Politics and economics in everyday life." In K. Monroe (ed.), The Political Process and Economic Change. New York: Agathon.
Kinder, D. R., S. J. Rosenstone, and J. M. Hansen
1983 "Group economic well-being and political choice." Pilot study report to the 1984 NES planning committee and NES board.
Kramer, G. H.
1971 "Short-term fluctuations in U.S. voting behavior, 1896–1964." American Political Science Review 65:131–43.
1983 "The ecological fallacy revisited: aggregate- versus individual-level findings on economics and elections and sociotropic voting." American Political Science Review 77:92–111.
Mebane, W. R., Jr.
1982 "The warp of sociotropic thinking." Paper delivered at the annual meeting of the Midwest Political Science Association, May 1982, Milwaukee, Wisconsin.
Neter, J., and J. Waksberg
1964 "A study of response errors in expenditures data from household interviews." Journal of the American Statistical Association 59:18–55.
Penick, B. K. E.
1976 Surveying Crime. Washington, DC: National Academy of Sciences.
Rosenstone, S. J.
1982 "Economic adversity and voter turnout." American Journal of Political Science 26:25–46.
1983 Forecasting Presidential Elections. New Haven: Yale University Press.
Rosenstone, S. J., J. M. Hansen, and D. R. Kinder
1983 "Measuring personal economic well-being." Report submitted to the board of overseers, National Election Study, October 1983.
Schlozman, K. L., and S. Verba
1979 Injury to Insult. Cambridge: Harvard University Press.
Sears, D. O., and R. R. Lau
1983 "Inducing apparently self-interested political preferences." American Journal of Political Science 27:223–52.
Sears, D. O., R. R. Lau, T. R. Tyler, and M. A. Harris, Jr.
1980 "Self-interest versus symbolic politics in policy attitudes and presidential voting." American Political Science Review 74:670–84.
Tufte, E. R.
1978 Political Control of the Economy. Princeton: Princeton University Press.

Race Differences in Abortion Attitudes

ELAINE J. HALL AND MYRA MARX FERREE

SINCE 1965 a consistent finding in public opinion surveys of abortion attitudes has been that black respondents are less in favor of legal abortion than white respondents. This is true whether one looks at the NORC General Social Surveys (Arney and Trescher, 1976; Granberg and Granberg, 1980; Evers and McGee, 1977) or at other surveys using differently worded questions (Westoff and others, 1969, using NFS; Lipson and Wolman, 1972, using a Population Commission survey; Pomeroy and Landman, 1972, using Gallup Polls). Only Combs and Welch (1982) take the further step and focus upon the variables which may explain this racial difference in prochoice attitudes.

Using the NORC surveys for 1972 to 1980, Combs and Welch attempt to explain racial differences in legal abortion support by three antecedents: (1) lower SES, as measured by family income and years of education, (2) stronger religious practice, as measured by frequency of church attendance, denominational membership (Catholicism and

---

Abstract Public opinion surveys since 1965 find that black respondents are less in favor of legal abortion than white respondents. Using the 1982 NORC General Social Survey, we replicate and expand one of the few studies (Combs and Welch, 1982) that examined the structure and determinants of prochoice attitudes of blacks and whites. Our major findings are (1) the racial difference in prochoice attitude is as great in 1982 as in the 1970s, (2) contrary to the suggestion of Combs and Welch, the demographic and attitudinal determinants of abortion attitudes differ for blacks and for whites, and (3) for those respondents who differentiate their acceptance of legal abortion, the pattern of prochoice attitudes also differs by race.

Elaine J. Hall is a doctoral student in the Department of Sociology, University of Connecticut, Storrs, and Myra Marx Ferree is an Associate Professor of Sociology in that department. The authors wish to thank Mark Abrahamson, Nancy Andes, and the anonymous reviewers for this journal for their constructive criticism of an earlier draft. The article was originally presented as a paper at the Eastern Sociological Society Annual Meetings, March 1985.

Public Opinion Quarterly Vol. 50:193–207 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-193/$2.50

fundamentalism), and the respondent's self-identified strength of affiliation, and (3) southern culture, measured by present region of residence and at age 16, and by present size of community and at age 16. Each of these variables has been shown to have a negative zero-order correlation with abortion support. Each can also be assumed to be more prevalent among blacks than whites. Therefore, Combs and Welch predict that controlling for these three antecedents would reduce or eliminate the racial difference in abortion attitudes.

In fact, Combs and Welch find that controlling for demographic variables (SES, southern culture, age and sex) reduces the difference between blacks and whites by only about one-third. When religious practice variables are also controlled, the black-white difference declines further. However, even when controlling for all 11 indicators, a small but statistically significant racial difference in abortion attitudes remains.

Although there is a statistically significant racial difference in each subperiod which Combs and Welch examine, the size of this difference appears to diminish over the 1970 decade. Combs and Welch therefore suggest that, with a few minor exceptions, "patterns in attitudes towards abortion in the black community mirror those in the larger society, and thus seem to be shaped largely by the same forces as white attitudes" (1982:517). In short, they see the differences between blacks and whites as essentially residual effects of historically different social positions that are converging over time.

Operating from a demographic perspective on abortion attitudes, Combs and Welch assume a longitudinal trend toward increased support of legal abortion in all groups as they become more urban, educated, and secular. These demographic trends are generally seen as producing more "modern" or progressive attitudes on a variety of attitudinal measures, of which acceptance of legal abortion is only one. In contrast, a feminist perspective assumes a conflict model of increasing polarization of attitudes between groups as the debate over women's equality and the family is politicized (Deitch, 1982; Mueller, 1983; Conover and Gray, 1983). This feminist explanation stresses attitudinal variables, measures of an underlying belief system regarding family and women's roles, as antecedents of abortion attitudes. From this perspective, availability of legal abortion is a central issue in obtaining reproductive freedom, and reproductive self-determination is a central issue in obtaining social, political, and economic equality for women (Luker, 1984).

Attitudinal variables suggested by the feminist perspective have been shown to be important predictors of abortion attitudes, even when demographic antecedents are controlled (Blake and DelPinal, 1980). Granberg and Granberg find that while demographic variables explain 10 percent of the variation in prochoice support in the 1972–80 NORC data, a personal morality index explains 21 percent of the variation. Attitudes regarding premarital sexual relations and ideal family size are the two strongest and more consistent predictors of abortion attitudes in this index (Granberg and Granberg, 1980:255–56). Deitch (1982) uses opinions on premarital sex, homosexuality, women's roles, liberal-conservative self-identification, and government spending priorities in an analysis of the 1972–80 GSS. She finds the associations between prochoice position and each of these attitudes do not disappear when education, sex, and religious affiliation are controlled. Weak positive correlations with prochoice are reported for both support for ERA (Granberg and Granberg, 1980:259) and for liberal-conservative attitudes toward economic and political roles for women (Granberg, 1978:422).

While the connection between position on legal abortion and other attitudes is important, it has never been explored separately for blacks. Other research on black attitudes toward feminism suggests that the determinants may be different for the two racial groups (Ransford and Miller, 1983). They find middle-class blacks more conservative in their attitudes towards women's roles than other blacks. This relationship is the reverse of that typically found for whites, and one that is not replicated on other non-sex-role issues, Their work also implies a different effect of gender in the two groups. Black men are found to be more traditional than white men, while black and white women are equally feminist when the demographic differences between them are controlled. Rather than the convergence anticipated by Combs and Welch's demographic model, the feminist perspective anticipates continued and even increased differences between blacks and whites if the demographic and attitudinal roots of a prochoice position are different in the two groups, as Ransford and Miller find the correlates of feminist attitudes to be. Moreover, Combs and Welch note that black respondents' attitudes toward abortion are less scalable than those of whites. This differential scalability may reflect differences of attitude structure in the two population groups that deserve further scrutiny.

The purposes of this article are to extend the analysis of Combs and Welch forward to the 1984 General Social Survey, to specify additional attitudinal determinants of the relative lack of support for legal abortion among blacks, and to compare the structure of white and black abortion attitudes. Specifically, we hypothesize the following:

1. The racial difference in abortion attitudes is persistent over time. We expect no evidence of a trend toward convergence when black and white attitudes are examined for the entire 1972–1984 GSS series.

2. The racial difference in prochoice attitudes is not due solely to positional antecedents. We expect a significant racial difference in prochoice attitudes to remain even when controlling for the demographic and religious practice variables examined by Combs and Welch.

3. The inclusion of attitudinal variables will enhance the explanatory power of the model but will not completely eliminate the difference by race. We expect both feminist and sexual morality attitudes to significantly increase the variance explained but to leave a statistically significant racial gap, expressing the political dimension of difference.

4. Because of the political component in prochoice position, we expect black and white attitudes to rest in part on different antecedents. This would be evident if an antecedent (such as gender) is significant for one race but not for the other, or if an antecedent (such as age) is significant for both races but varies in strength or direction of influence. Among these interactions, we expect feminist attitudes to be more predictive of a prochoice position for whites than for blacks, who may have conflicting political priorities.

5. The structure of abortion attitudes is itself different for blacks and whites. When support for a prochoice position is qualified by situation, we expect to find different patterns of partial support for blacks and whites.

Methods

Because of the small number of black respondents in each yearly NORC survey from 1972 to 1980, Combs and Welch combined their data into periods of three years each: 1972–74 (671 blacks), 1975–77 (468 blacks), and 1978–80 (298 blacks). Pooling data hides variation within each period; this plus the small number of black respondents in the last of these combined GSS samples makes it difficult to determine trends.[1] We expand the series to 1984 and break out each year of the GSS for trend analysis, but then concentrate on the 1982 GSS, where the oversampling of blacks eliminates the need to pool data to get an adequate number of cases for analysis of black attitudes. After excluding 27 respondents of "other" races, the total 1982 sample of 1833 consists of 28 percent (510) blacks and 72 percent (1323) whites. When black and white respondents are analyzed together, weighting is used to make the sample representative (Smith, 1984).

---

[1] The interpretation of trends across the three time periods is further complicated by sampling issues. The 1972–74 surveys were a modified probability design which used quotas for age, sex, and employment status on the block level. Because the 1972 GSS was drawn on the 1960 census information, the proportion of blacks in this survey is atypically high (261, or 16.2 percent of the 1613 total). In the 1975–77 period, problems of refusal and completion rates produced a low proportion of blacks in the 1976 GSS (only 129 or 8.6 percent of the 1499 total). Although a full probability sampling design was used in the 1978–80 period, the total number of respondents is limited to only two surveys (1978 and 1980). Overall, these sampling issues combine to give the effect of a diminishing number of black respondents in the three periods defined by Combs and Welch.

Following Combs and Welch, we construct a six-point additive scale of prochoice attitude by summing positive responses to each of the six questions about situations where abortion should or should not be legal. The demographic predictors are measured as follows: (1) SES by years of education and family income,[2] (2) gender and age of the respondent, (3) Southern culture by region presently and at age 16 (1 = South, 2 = non-South), and by size of community presently (1 = city, 2 = suburb, 3 = other) and at age 16 (1 = nonrural, 2 = rural), and (4) religious practice variables measured by frequency of church attendance, strength of affiliation, and dummy variables for Catholicism and fundamentalism. For the following attitudinal variables, the smaller the value, the more traditional the definition of women's roles: (1) extent of support for ERA, (2) attitude toward premarital sexual relations, (3) belief that women should stay home and leave running the country to men, and (4) belief that married women should not work. Attitude toward family size reflects the centrality of motherhood in the definition of women's roles; the larger the ideal family size, the more traditional the definition of women's roles.

Results

THE TREND ANALYSIS

First, we examine the racial difference in prochoice attitudes from 1972 to 1984 for both the additive prochoice scale and for each of the six reasons for an abortion. In Table 1 we note a substantial and statistically significant difference in prochoice attitudes between blacks and whites for each year. The mean scale scores reveal that whites are more prochoice for each of the 11 years, a difference that averages approximately three-quarters of a point on the six-point scale. Whites are also consistently more prochoice than blacks for each of the six reasons, although the difference is not significant for each reason for each year. The mean percentage difference per reason ranges from

---

[2] Income has been recoded to the category midpoint in dollars for each year. The income variable used in each survey was the one designed to reflect the income distribution at that point; with inflation, categories have changed. For details, see pages 48–51 in Davis and Smith (1982). In all other cases, the coding follows Combs and Welch (1982:512–13).

Table 1. Distribution of Abortion Support by Race, 1972–1984

	Health			*Rape*			*Defect*			*Can't Afford*			*Single*			*Wants No More*			*Prochoice Scale*		
Year	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*W*	*B*	*Diff*	*White*	*Black*	*Diff*
1972	86	70	16*	78	56	22*	78	56	22*	46	36	10*	43	32	11*	39	30	9*	3.71	2.76	.95*
1973	92	80	12*	83	66	17*	85	67	18*	54	38	16*	50	30	20*	48	36	12*	4.10	3.15	.95*
1974	91	87	4	84	71	13*	84	71	13*	54	42	12*	50	33	17*	46	36	10*	4.09	3.40	.69*
1975	89	82	7*	82	68	14*	82	63	19*	51	45	6	47	37	10*	44	40	4*	3.95	3.38	.57*
1976	90	81	9*	82	67	15*	83	65	18*	52	41	11*	50	33	17*	45	38	7	4.01	3.22	.89*
1977	90	77	13*	83	65	18*	86	67	19*	53	43	10*	49	35	14*	45	42	3	4.04	3.22	.82*
1978	89	85	4*	83	65	18*	82	67	15*	46	40	6	40	33	7	39	36	4	3.78	3.25	.53*
1980	89	81	8*	82	70	12*	82	66	16*	50	53	7	48	36	12*	46	40	6	3.96	3.34	.62*
1982	90	82	8*	85	68	17*	83	66	17*	50	39	11*	48	31	17*	47	34	13*	4.02	3.19	.83*
1983	88	82	6*	80	72	8*	78	66	12*	42	39	3	39	26	13*	38	35	3*	3.58	3.15	.43*
1984	89	81	8*	80	63	17*	80	63	17*	46	35	11*	45	30	15*	43	30	13*	3.80	3.01	.79*
Mean percentage diff			8.64*			15.55*			16.91*			9.36*			13.91*			7.55*			
Mean item score	1.11	1.20		1.18	1.34		1.18	1.35		1.50	1.60		1.54	1.68		1.56	1.64		3.91	3.17	.74*
Race-year interaction (*F*)	2.58 ($p < .005$)			1.35 ($p > .20$)			.82 ($p > .61$)			.72 ($p > .71$)			.88 ($p > .55$)			1.03 ($p > .55$)			1.19 ($p > .29$)		

* Significant at .05 level.

almost 9 to 17 percentage points. The smallest race difference is found in the item where whites are least supportive: when the woman is married and wants no more children.

Second, the linear trend of diminishing differences reported by Combs and Welch is not supported in these data. Despite fluctuations in the attitudes of both whites and blacks throughout the period, the most meaningful change is an increase in support by both groups between 1972 and 1973, probably reflecting the Supreme Court's legalization of abortion at that time. There is no consistent trend toward convergence between the two groups for either the individual reasons or the prochoice scale for the full 1972–84 period or for the more restricted set of 1972–80 data that Combs and Welch used. For both of these periods, the *F* statistics for the year–race interaction are not significant for either the prochoice scale nor, with one exception, for the individual reasons.[3] Looking at abortion when the mother's health is in danger, the one item that shows significant convergence, we find that with the exception of 1972, the white marginals can be represented as 90 ± 2 and those of blacks as 82 ± 5. The interaction thus seems to show that in aftermath of *Roe v. Wade* on this item, blacks increased their support more than whites, who were already very supportive. For the other items, whites as well as blacks increased their approval, so the magnitude of the difference did not decline.

Looking at the prochoice scale in Table 1, we can see how pooling data into three periods (1972–74, 1975–77, and 1978–80), as Combs and Welch did, could create the appearance of a trend toward disappearing differences. The years of least difference (1978 and 1983), however, are notable for their atypically low levels of white approval of legal abortion for the more discretionary (often called "soft") reasons, rather than for increases in black support. Overall, we can conclude that the stability of abortion attitude is much more impressive than any changes for either blacks or whites.

EXPLAINING THE RACE GAP

Turning now to attempts to explain this persistent race difference, we see in Table 2 that the demographic and religious practice variables employed by Combs and Welch continue to have about the same effect on the impact of race on abortion attitudes as they did in earlier years. In 1982, the unstandardized regression coefficient (−.83) for race is substantially reduced (to −.71) with the control of the demographic variables, and is reduced even further (to −.55) but remains significant

---

[3] Combs and Welch also note that the trend toward convergence in their data is not significant using a chi-square goodness of fit test (Taylor, 1980).

Table 2. Impact of Race on Abortion Attitudes, 1972–1984[a]

	1972–1974			*1975–1977*			*1978–1980*			*1982*			*1983–1984*		
Variables	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2	*b*	*(beta)*	R^2
Race alone	−.90*	(−.15)	.02	−.70*	(−.11)	.01	−.57*	(−.09)	.01	−.83*	(−.13)	.02	−.60*	(−.09)	.01
Race with controls for demographic variables[b]	−.49*	(−.08)	.12	−.43*	(−.07)	.09	−.32*	(−.05)	.08	−.71*	(−.11)	.13	−.42*	(−.07)	.08
Race with controls for demographic & religious practice variables[c]	−.45*	(−.08)	.24	−.27*	(−.04)	.23	−.20	(−.03)	.23	−.55*	(−.09)	.25	−.24	(−.04)	.22
N[d]		4054			4019			2659			1604			2627	

* Significant at .05 level.

[a] Since this article does not intend to be an exact replication of Combs and Welch, we code some variables differently than they did. We use XNORCSIZ instead of SRCBELT to measure present size of community; either variable is coded 1 = city, 2 = suburb, and 3 = other. To allow for inflation, four income variables are used to cover the 1972–1984 period (INCOME72, INCOME, INCOME77, and INCOME82). All income variables are coded to the midpoint of the category. Frequency of church attendance is coded to number of weeks per year to make it a legitimately interval variable.

[b] Demographic variables: education, income, age, sex, region, size of community, region and size of community at age 16.

[c] Religious practice variables: attendance, Catholicism, fundamentalism, strength of affiliation. Strength of affiliation is not available for the 1972–1974 period and therefore is not included in the regression for that period.

[d] These results are based on pairwise deletion. The regressions were also conducted using listwise deletion, revealing no substantial differences. The 1982 regression was weighted to correct for the oversample of blacks.

when religious practice variables are controlled. The size of the residual gap is thus roughly half a point on a six-point scale. The same pattern occurs for the 1983–84 period (a coefficient of −.60 is reduced to −.42 and further to −.24) except that the racial difference with all variables in the regression is no longer significant, contrary to our hypothesis. It is difficult to interpret this as a trend, however, since the 1983 difference is at least as large as the difference in 1972–74. The substantial amount of variation explained by demographic and religious practice variables during the 1970s (adjusted R^2 of .24, .23, .23 for the three pooled time periods) continues in the 1980s (adjusted R^2 of .25 and .22).

In the years for which these measures are available, the inclusion of attitudinal variables as antecedents to abortion attitudes further enhances the explanatory power of the model but does not diminish the racial differences in prochoice attitudes. Four attitudinal variables—premarital sex, ideal family size, women working, and women staying out of politics—were asked in the 1975, 1977, and 1982 General Social Surveys. Adding them as a fourth step to the regression used in Table 2 explains an additional amount of the variance in abortion attitudes. The adjusted R^2 increases from .21 to .28 in 1975, from .24 to .32 in 1977, and from .25 to .34 in 1982. In all three years when this set of attitude variables is available, controlling for feminism and fertility attitudes actually increases the residual difference by race. The size of the gap varies from −.25 in 1975 to −.46 in 1977 and −.66 in 1982, thus showing a slight increase over this period.

Further examination of the antecedents of black attitudes toward

Table 3. The Impact of Demographic and Attitudinal Variables on Abortion Attitudes for Blacks and for Whites, 1982

	Blacks					*Whites*					*F* (race inter-action)
		Demographic		*Attitudinal*			*Demographic*		*Attitudinal*		
	Simple r	*b*	*(Beta)*	*b*	*(Beta)*	*Simple r*	*b*	*(Beta)*	*b*	*(Beta)*	
Income	.21*	.013	(.08)	.013	(.08)	.19*	.012*	(.08)	.010*	(.07)	0.96
Education	.32*	.085*	(.14)	.059	(.10)	.26*	.114*	(.18)	.091*	(.14)	0.20
Sex	.00	.364	(.09)	.425*	(.10)	−.03	.101	(.03)	.178	(.05)	1.44
Age	−.29* +	−.013*	(−.12)	−.012* +	(−.10)	−.11* +	.004	(.04)	.016* +	(.15)	10.08*
Size of residence	−.19*	−.169	(−.08)	−.144	(−.06)	−.15*	−.175*	(−.08)	−.128*	(−.05)	0.07
Region	.24* +	.101	(.02)	−.115	(−.00)	.10* +	−.147	(−.03)	−.228	(−.05)	0.89
Rural residence 16	−.29*	−.526*	(.12)	−.391	(.09)	−.20*	−.312*	(.07)	−.155	(.04)	0.42
Region 16	.30* +	.458	(.10)	.486*	(.11)	.11* +	.125	(.03)	.016	(.00)	1.34
Catholicism	.04	−.138	(−.02)	−.159	(−.02)	−.06*	−.384*	(−.09)	−.382*	(−.08)	0.17
Fundamentalism	−.10* +	−.002	(−.00)	.083	(.02)	−.22* +	−.381*	(−.08)	−.142	(−.03)	0.30
Attendance	−.26* +	−.018*	(−.22)	−.009*	(−.10)	−.38* +	−.025*	(−.34)	−.014*	(−.19)	0.51
Premarital sex	.37* +			.352* +	(.23)	.48* +			.529* +	(.34)	4.90*
Ideal family size	−.26*			−.121*	(−.12)	−.23*			−.177*	(−.13)	0.21
Pro-ERA	.10* +			.252*	(.11)	.23* +			.215*	(.10)	0.71
Female home	.28*			.304	(.07)	.23*			.106	(.02)	0.01
Female work	.05			.021	(.01)	.01			−.036	(−.02)	0.05
Adjusted R^2		.22		.29			.23		.34		
Constant		1.13		−0.62			2.73		0.87		

NOTE: Unstandardized regression coefficients are given first with betas in parentheses.
* Individual coefficient significant at .05 level.
+ Differences in *r* or *b* between blacks and whites are significant at .05.

abortion requires focusing on the 1982 GSS, where the oversample of blacks gives sufficient cases for reliable estimates. As evident in Table 3, the inclusion of the four attitudes mentioned above and a further measure of feminism, support for the ERA, substantially increases the amount of variance explained in the abortion attitudes of both blacks (from an adjusted R^2 of .22 to .29) and whites (from an adjusted R^2 of .23 to .34). In fact, attitude toward premarital sexual relations becomes the best predictor of abortion attitudes for both blacks and whites.

In Table 3, we present the comparison by race of the impact of the demographic-religious variables suggested by Combs and Welch and our attitudinal variables. Reviewing the zero-order correlations (columns 1 and 6) shows seven variables which are differently associated with prochoice attitudes for blacks and for whites. The negative association of church attendance with prochoice attitudes is significantly stronger for whites than for blacks. Respondents of both races who have a sexually permissive attitude, are not fundamentalist, or are pro-ERA are more likely to be prochoice, but these are also significantly stronger relationships for whites than for blacks. The relationship of age with abortion attitudes (more prochoice if younger) and of region (current residence and region at age 16—more prochoice if from the North) is stronger for blacks than for whites. All seven variables continue to have somewhat different effects by race in the final regression, but only a few of these differences are statistically significant.

We test the significance of differences by race in the effects of our predictor variables by including race and a separate term for the interaction of race with each variable in an equation for the entire sample. This regression shows when effects are significantly different in magnitude or direction for blacks and whites. The *F*-tests for these interactions are given in column 11 of Table 3. We find only two significant differences: attitudes toward premarital sexual relations and age. The nature of these differences is most readily visible by examining the coefficients of regressions run separately by race. As we see in columns 4 and 9 of Table 3, which report these separate regressions, attitude toward premarital sexual relations, while significant for both groups, is considerably more important for whites. Age affects both blacks ($b = -.012$) and whites ($b = +.016$), but in opposite directions. For blacks, all other things being equal, younger respondents are more likely to be prochoice, while for whites, older respondents are more likely to be prochoice once factors such as fundamentalism and feminism are controlled.

Another type of difference evident in the paired regressions consists of those variables which are significant for one race but not for the other. If these differences were purely a function of sample size, we would expect all these variables to be significant only for the whites. Instead, we find a pattern of four variables which are significant only for whites and two variables which are significant only for blacks. Both measures of social class—income and education—are independently significant for whites but neither of them is for blacks. As would be expected, Catholicism is significant for only white respondents. Urbanness of present community is significant for whites but not for blacks, while region at age 16 is significant for blacks but not for whites. That is, whites presently living in rural areas are less prochoice, but for blacks, it is not the ruralness of their current community but growing up in the South that means they are less prochoice in abortion attitudes. We interpret this to mean that blacks tend to carry some normative aspects of their rural upbringing with them that influence their abortion attitudes even if they are presently living in the urban North, while whites who move to more rural or urban areas are more likely to conform to the norms prevalent there. Women are not more prochoice than men in the zero-order relationships, but a sex difference emerges once other factors are controlled. The positive effect of being female is strong enough to be significant for blacks but not for whites. Since this difference emerges even when other feminist attitudes are controlled, it is consistent with but not explained by the greater traditionality of black men that Ransford and Miller found in other sex role attitudes (1983).

DIFFERENCES IN ATTITUDE STRUCTURE

Having specified some differences in the determinants of abortion attitudes for blacks and whites, we turn now to differences in the structure of the attitude itself. Previous studies have typically either defined abortion attitudes in terms of the degree of total acceptance for all reasons (as we do, following Combs and Welch, above), or in terms of a dichotomy of the more generally accepted "hard" reasons of mother's health, rape, or birth defect in the fetus, and the less popularly supported "soft" reasons that the mother is single, can't afford, or wants no more children. Either of these strategies is useful for distinguishing the absolutists: those respondents who accept abortion for all six reasons (21 percent of all the blacks and 39 percent of all the whites) and those who oppose abortion for all six reasons (15 percent of all the blacks and 8 percent of all the whites). However, using either of these cutting points obscures the responses of those respondents who have a mixture of pro- and anti-abortion attitudes.

To clarify the structure of attitudes among the nonabsolutist blacks and whites, we employ a scale developed by Deitch (1982). Those who accept all the "hard" reasons but not all the "soft" reasons are classified as having mild anti-abortion attitudes (33 percent of the blacks and 39 percent of the whites). Those who oppose one or more of the "hard" reasons but accept at least one reason overall are classified as having strong anti-abortion attitudes (32 percent of the blacks and 15 percent of the whites). Table 4 presents the percentages accepting each of the individual reasons for those who have mild and strong anti-abortion attitudes.

Among respondents with a mild anti-abortion attitude, blacks and whites show a different pattern of acceptance of the "soft" reasons. Whites make few distinctions between the three "soft" reasons—their acceptance varies only 6 percent between "wants no more" (18 percent) and "can't afford" (24 percent). On the other hand, blacks make more distinctions among the social reasons they accept as legitimate—their acceptance varies 21 percent between the three reasons (only 13 percent accept abortion for a single woman but 34 percent want abortion to be legal for a woman who can't afford another child).

For whites with a strong anti-abortion attitude, there appear to be three distinct levels of acceptance of the individual reasons: woman's health, the other two "hard" reasons of rape and defect, and the three "soft" reasons which each receive less than 10 percent acceptance. Although blacks with a strong anti-abortion attitude share the major break in support between woman's health and the other two "hard" reasons, the distinction between the "hard" and "soft" reasons is less

Table 4. Percentage Distribution of Individual Reasons Accepted by Mild and Strong Anti-abortion Respondents by Race, 1982

	Mild Anti-abortion[a]				*Strong Anti-abortion*[a]			
	Black		*White*		*Black*		*White*	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Woman's health	100	na	100	na	88	12	83	17
Rape	100	na	100	na	44	56	48	52
Birth defect	100	na	100	na	37	63	38	62
Can't afford	34	66	24	76	22	78	9	91
Single	13	86	21	79	17	83	8	92
Wants no more	24	76	18	82	15	85	9	91
% of race in this category	33		39		32		15	
N	170		516		156		193	

[a] Mild anti-abortion attitude consists by definition of those who say "yes" to all three "hard" reasons (health, rape, defect) and "no" or "don't know" to one or more of the "soft" reasons. Strong anti-abortion attitude consist of those who say "no" to one or more of the "hard" reasons but did not answer "no" to all reasons.

pronounced. Blacks are more likely than whites to accept some of the social reasons even when they have a strong anti-abortion attitude. In addition, blacks again show a pattern of more differentiation between the three social reasons: whites vary in their acceptance between the three social reasons only by 1 percent, while blacks vary by 7 percent (15 percent want abortion legal for women who want no more children but 22 percent do for women who can't afford another child). It appears that among blacks, social reasons, especially economic need, are perceived as "harder" or more legitimate than whites see them. Moreover, the nonabsolutist blacks are consistently differentiating more between the three "soft" reasons they accept as legitimate than are whites.

Summary and Conclusions

All five hypotheses of this article are at least partially supported. First, new and better data on blacks have made it possible to see that the apparent convergence between black and white abortion attitudes in the period 1972–80 is misleading. The gap in prochoice attitude is as great in the 1982–84 period as it was 10 years before, and the pattern of lesser black support for legal abortion is consistent on an item-by-item basis over the decade.

Known differences in the characteristics of the black and white populations (in SES, religiosity, and Southern culture) explain some but not all of this persistent difference. Contrary to our hypothesis, the race gap that remains once these factors are controlled is not significant in 1983–84. The race difference is still significant in 1982. In those years where data are available, controlling directly for attitudes toward women and sexuality increases the explanatory value of the model but actually increases the racial gap between black and white abortion attitudes. These attitudinal variables are not available for the years of least difference. Nonetheless, it appears that the difference between the races is not solely attributable to their different social positions and attitudes.

Looking more closely at 1982, the pattern of determinants of abortion attitudes is found to differ for blacks and whites. Education and income, rural residence, and Catholicism are significant for whites but not blacks, while gender and Southern upbringing contribute significantly to the explanation of blacks' attitudes but not whites'. Attitude toward premarital sex is significant for both blacks and whites, but significantly stronger for whites, while the effects of age are equal and opposite in the two groups. Contrary to our fourth hypothesis, however, the link between abortion and specifically feminist attitudes is not significantly different for blacks than for whites.

For those respondents who differentiate their acceptance of legal abortion by situation, the pattern of prochoice attitudes differs by race as predicted. Black respondents are more likely to accept the social reason "married and can't afford more children" as legitimate than whites are, and more likely to accept this reason than that the woman is single and does not want to marry the father. White respondents who differentiate situationally are about as likely to accept being single as being poor as a legitimate reason and draw fewer distinctions among the social reasons in general.

One surprising difference between the two races is in the effect of age. With these demographic and attitudinal factors controlled, it is the older whites who are prochoice, while it is younger blacks who, all other things being equal, accept legal abortion. Since age is typically related to more conservative attitudes, the residual positive effect of age for whites is somewhat paradoxical. One possibility is that, among whites, there might be a small conservative component in support for legal abortion that emerges when the more obvious conservative factors (fundamentalism and antifeminism) are controlled. This conservative component may be an expression of symbolic racism (Sears and others, 1979), that is, seeing legal abortion as an economic alternative to supporting a supposedly black welfare population. The absence of such an antiwelfare component in black prochoice attitudes might explain both the differing effect of age and some of the lesser support for legal abortion among blacks. Another possibility is a cohort effect in which older whites, but not older blacks, have more experience with illegal abortion and resist return to the danger and social stigma of that period. But this speculation deserves empirical investigation that is beyond the scope of the present article.

Given the differences in black and white prochoice attitudes found here, exploring the political bases of these positions seems a promising direction for further research. Luker's insightful study of prochoice and prolife activists reveals how abortion attitudes are intricately interwoven with beliefs about motherhood, children, female roles, sexuality, and morality. Since each side is defending a world view that it does not and can not question, the heated ideological nature of the debate is understandable. The particular attitudes involved may vary by subculture. Some authors, including Combs and Welch, refer to the abortion-as-genocide theme as a possible explanation of black anti-abortion attitudes. We think this is possible, but that this is not the major explanation of subcultural differences.

Our findings suggest that blacks, like whites, connect legal abortion to other aspects of women's rights, but that the pattern of support varies. Attitudes toward premarital sex are not as important for blacks, and the situation of a single woman who does not want to marry the father is also perceived as a less acceptable reason for abortion than is economic need. Control over women's sexuality thus seems less central to blacks' attitudes toward abortion. We suggest the explanation lies in the definition of women and family values found in black culture. Black women are seen as crucial support for family and community, and women are considered primarily in terms of the home, but the home is not particularly defined by marriage. Family values are important, but families need not be male-headed to be legitimate.

In sum, white as well as black attitudes toward legal abortion seem to be based on cultural and political factors that are largely similar but include some important differences by race. After controlling for demographic, religious practice, and attitudinal variables, we are left with an irreducible difference between blacks and whites. Black anti-abortion attitudes, like white views, reflect traditional values of their subculture. For blacks this may include accepting single women as adequate mothers. For whites, there may be an element of symbolic racism. For both, the role of personal experience in the formation of prochoice attitudes remains to be investigated.

References

Arney, William, and William Trescher
1976 "Trends in attitudes toward abortion, 1972–1975." Family Planning Perspectives 8 (3):117–24.

Blake, Judith, and Jorge H. DelPinal
1980 "Predicting polar attitudes toward abortion in the United States." Pages 29–56 in James Burtchaell (ed.), Abortion Parley. Kansas City: Andrews and McNeel.

Conover, Pamela Johnston, and Virginia Gray
1983 Feminism and the New Right: Conflict over the American Family. New York: Praeger Publishers.

Combs, Michael, and Susan Welch
1982 "Blacks, whites, and attitudes toward abortion." Public Opinion Quarterly 46 (4):510–20.

Davis, James A., and Tom W. Smith
1982 General Social Surveys, 1972–1982: Cumulative Codebook (NORC edition). Roper Public Opinion Research Center, University of Connecticut, Storrs.

Deitch, Cynthia
1982 "The new right, feminism, and abortion: patterns of public opinion, 1972–1980." Paper presented at the American Sociological Association meetings.

Evers, Mark, and Jeanne McGee
1977 "The trend and pattern in attitudes towards abortion: 1965–1976." Paper presented at the American Sociological Association Meetings, Chicago.

Granberg, Donald
1978 "Prolife or reflection of conservative ideology? an analysis of opposition to legalized abortion." Sociology and Social Research 62 (3):414–29.

Granberg, Donald, and Beth Granberg
1980 "Abortion attitudes, 1965–1980: trends and determinants." Family Planning Perspective 12:250–61.

Lipson, Gerald, and Dianne Wolman
1972 "Polling Americans on birth control and population." Family Planning Perspectives 4:39–42.

Luker, Kristin
1984 Abortion and the Politics of Motherhood. Berkeley: University of California Press.

Mueller, Carol
1983 "In search of a constituency for the 'new religious right.' " Public Opinion Quarterly 47:213–29.

Pomeroy, Richard, and Lynn Landman
1972 "Public opinion trends: elective abortion and birth control services to teenagers." Family Planning Perspectives 4:44–55.

Ransford, H. Edward, and Jon Miller
1983 "Race, sex and feminist outlooks." American Sociological Review 48 (1):46–59.

Sears, David O., Carl P. Hensler, and Leslie K. Spier
1979 "Whites' opposition to 'busing': self-interest or symbolic politics." American Political Science Review 73 (2):369–84.

Smith, A. Wade
1984 "Evaluating the products of alternative sampling methods." Paper presented at the American Sociological Association Meetings, San Antonio.

Taylor, D. Garth
1980 "Procedures for evaluating trends in public opinion." Public Opinion Quarterly 44:86–100.

Westoff, Charles, Emily Moore, and Norman Ryder
1969 "The structure of attitudes toward abortion." The Milbank Memorial Fund Quarterly 47:11–36.

Picking the Winners: Politician vs. Voter Predictions of Two Controversial Ballot Measures

JAMES B. LEMERT

PUBLIC opinion researchers have known for some time that members of the public often can be wildly incorrect when estimating how other members of the public feel about issues. This general finding has been called *pluralistic ignorance* by some (e.g., O'Gorman and Garry, 1976). Roughly at the same time, students of legislative representation such as Miller and Stokes (1963) were developing evidence—disputed as an artifact in some quarters—implying that legislators' perceptions of their constituents' opinions were more accurate as issue salience/ controversy increased.[1]

Justifications of representative democracy implicitly assume that politicians are actually capable of perceiving the policy wishes of their constituents. Obviously, any model of representation that requires

---

[1] See especially Achen (1977), who asserts that use of the correlation coefficient inevitably meant that members of the 1958 House of Representatives would show greater "accuracy" in their estimates of constituency opinion on civil rights, where between-district variation was greatest, than on social welfare and foreign policy, where regional differences were not as great. In the present study, both measures studied were controversial and both displayed marked differences in support between areas of the state. However, unlike Miller-Stokes and similar "representation" studies, we asked for predictions of *statewide* results, not of results in any district(s).

Abstract Both politicians and voters were asked to predict outcomes of two Oregon ballot measures in 1982. As expected, politicians' predictions always were closer to the mark than voters' were. Further, voters showed stronger signs of wishful thinking (the "Looking-Glass effect") in their predictions than did politicians. Using published preelection polls apparently improved politicians' accuracy in 1982, as well as voters' accuracy in a separate 1984 survey. No other sources of data improved predictive accuracy. Findings have implications for theories of representative government and are consistent with a new theory of public opinion.

James Lemert is Professor and Director of the Division of Communication Research, School of Journalism, University of Oregon.

Public Opinion Quarterly Vol. 50:208–221 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-208/$2.50

elected politicians to consider public opinion in decisions assumes that these decision makers' perceptions of public opinion are accurate "enough." But what is "enough"? Though Hedlund and Friesema found that, in each case, Iowa legislators were right much more often than wrong in predicting how their constituents would vote on four ballot measures, these authors interpreted their results as raising serious questions about representative democracy (1972: 410–11).

A major problem with Hedlund and Friesema's study design is that it considered only predictions made by politicians, merely comparing various subsets of them against other subsets. This seems a tenuous basis on which to conclude that the entire theory of representative democracy is threatened. In the absence of a baseline that is external to the sample of elected politicians, it is difficult to say what percentage of accuracy is sufficient. Fortunately, however, the pluralistic ignorance literature suggests—almost by definition—that politicians should at least do better than the mass public.

Lemert (1981: 12) has based a key part of his public opinion theory on the as-yet-untested proposition that perceptions of public opinion held by politicians and other decision makers will be more accurate than perceptions held by ordinary citizens, as a result of a series of constraints experienced by politicians. Constraints may include such things as (1) questions from journalists about decision makers' public opinion perceptions and their evidence for them, (2) contrary perceptions offered by rival decision makers, (3) uncertainty as to whether polls or other independent sources of public opinion data might surface at some future time, and (4) at least for elected politicians and their advisors, the continuing opportunity (or threat) presented by the need to face the voters.

No such pressures are felt by ordinary members of the mass public (Lemert 1981: 13–15). Without these constraints, the result for ordinary members of the public can be what Fields and Schuman (1976: 445) call the Looking-Glass perception: "In the absence of strong counter-forces, a large proportion of people feel that the world they live in agrees with their own opinion on public issues."

A kind of pressure upon ordinary citizens has been proposed by Noelle-Neumann in her writings about the "Spiral of Silence." In her book she tends to assume a kind of "quasi-statistical" accuracy on the part of ordinary citizens, but even there, she notes, inaccuracy can be created through the Looking-Glass effect and through inaccurate media portrayals of dominant opinion (1984: 124–27; 157–64). And in an earlier journal article (1974: 45), she explicitly asserts that the spiral of silence itself can mislead both the vocal, active people and the silent ones about the true distribution of opinion.

In the present study, we treat the way people voted in an election as an indicator of their "true" attitude on that issue. (Their reported vote is obtained *before the outcome of the election is known.*) We then make two predictions: (1) politicians will be much more accurate predictors of outcomes for two highly controversial state ballot measures than will a comparison group of voters, and (2) whatever apparent signs of the Looking-Glass phenomenon occur will be stronger among voters than among politicians.

This study cannot directly test Lemert's argument that greater pressure or constraint makes politicians more accurate. Nevertheless, his argument predicts considerably higher accuracy among the politicians. If this prediction is fulfilled, then we can begin to consider alternative explanations for it.

Method

For at least two reasons, asking voters to predict the outcome of elections may be a fairer test of pluralistic ignorance than is the practice of asking a general survey sample to estimate survey responses given by others. First, general samples obviously can include respondents who lack acquaintance with the issue(s) being asked about, thus producing artifactual overestimates of the breadth and depth of pluralistic ignorance. In contrast, people who have just voted on an issue presumably have some acquaintance with it. Second, issues in elections have a kind of history and concreteness that issues asked about in surveys may not. For example, if one is in the habit of trying to predict election results, one finds out the actual result soon enough. But if one is trying to predict attitudes toward racial integration in Detroit, for example (Fields and Schuman, 1976), how does one find out whether one is right or wrong?

The principal comparisons will be made between politician and voter predictions of two ballot measure outcomes in November 1982. In addition, certain supplementary data will be used from a second exit poll of voters, conducted this time at the November 1984 election.

The 1982 election in Oregon featured two highly publicized and controversial state ballot measures: Measure 3 (a property tax limitation measure patterned after California's Proposition 13) and Measure 6, which would have eliminated statewide land-use planning controls. The 1984 election included another property tax limitation measure which was virtually identical to the 1982 proposal. The three measures had characteristics we thought were ideal for our purposes. First, there was considerable uncertainty about whether each would pass, because polls early in the campaigns had shown all three of them with large leads while more recent polls showed their leads eroding rapidly. Second, each issue had a considerable track record in Oregon, close variants of each having been defeated in the past. Third, each issue clearly was salient to significant numbers of voters, judging by how frequently and rapidly petitioners had placed each on the ballot.

THE VOTER SAMPLES

The 1982 and 1984 voter groups were produced by a multistage, stratified random sampling procedure. The first stage was to obtain a sample of Eugene-Springfield (OR) metropolitan area precincts (29 in 1982, 41 in 1984), stratified by relative wealth and past liberal-conservative voting records. Precincts were chosen randomly within each of these strata. The next stage involved a sample of voting times. On the basis of Lane County Election Department estimates of turnout by daypart, half the interviewers went to precincts between 5 P.M. and the closing of the polls at 8 P.M., a quarter went between 1 P.M. and 5 P.M., and a quarter between 8 A.M. and 1 P.M. Within this overall time allocation, dayparts were assigned randomly to precincts. Finally, each interviewer was given a procedure for selecting specific interview respondents. A "skip interval" number between one and five was randomly assigned to each precinct. Interviewers given the interval number three, for example, would have to wait for three voters to leave before interviewing the next person. Interviewers resumed counting after completing each interview. Each interviewer was to obtain 10 respondents during the assigned time period. The 1982 survey produced 283 completed interviews in 29 precincts; the 1984 survey produced 408 in 41.

THE SAMPLE OF POLITICIANS

Mail questionnaires were sent before the 1982 election to 115 prominent Oregon politicians, all but three of whom were themselves involved in a race for state or U.S. office. On the assumption that pressure toward accuracy would be greater when a candidate had an opponent, only candidates with opponents were sampled. Three prominent elected state executives who were not directly involved in the 1982 election also were included in the mail sample. Their actions and decisions often had been intimately associated with land-use planning and property tax issues, so they could be expected to be constrained toward accurate perceptions of public opinion.

Names and addresses of candidates were obtained from the Oregon Secretary of State's office. Assuming normal delivery of the mail, each questionnaire arrived at the listed address seven days before the election. All politicians were instructed to mail their responses back so that the envelope would be postmarked *before* election day. Of the 115 questionnaires sent out, 58 usable replies were received. Four replies were discarded because they were postmarked too late, two envelopes were returned as undeliverable, and three politicians sent notes saying their questionnaire was received too late. The effective completion rate, therefore, was 58 of 106, or 54.7 percent.

The most common office sought by respondents was in the State House of Representatives (35 replies), with the next most common (13 replies) being the five U.S. Congress seats plus three contested state executive offices. Slightly more than half the replies came from incumbents (or, in a few cases, their campaign managers).

THE PREDICTION QUESTIONS

Similar questions were asked voters (1984) and both voters and politicians (1982) about each ballot measure. Before any other questions, respondents were first asked to predict whether each measure would pass or fail. Then they were asked to estimate the percentage of "yes" votes the measure would get from Oregon voters.[2]

Results

Both the 1982 and 1984 property tax measures were defeated narrowly by Oregon voters. The 1982 measure received 49.5 percent of the

---

[2] Question wordings differed slightly because voters were questioned in person while politicians filled in a mail questionnaire. Here is how politicians were asked the questions:

> Please write in your best estimate, based on the information you presently have, about how Measure 3—the property tax limitation measure—will do, *statewide, in percentage terms*.
>
> I think Measure 3 will get ____ percent "*yes*" votes statewide and thus will (circle one) pass. fail.

The question for Measure 6 was identical except for the measure's number and the substitution of "the LCDC 'repeal' " for "the property tax limitation measure."

Interviewers asked the comparable questions of voters as follows:

> I'd like to ask you to try your hand at predicting whether Oregon voters will pass Measures 3 and 6. . . . First, let's do Measure 3, the property tax limitation measure. Do you think Oregon voters will approve Measure 3, or do you think it will fail, statewide? What percentage of the vote will the YES side get, then, given that you think it will . . . (pass) (fail)?

The questions for Measure 6 were similar.

In 1984, the questions to voters differed slightly from the 1982 voter questions by explicitly asking them to ignore their own wishes:

> As you know, there's been lots of predicting about how Measure 2, the property tax limitation measure, will do. I'd like you to take a stab at guessing whether Oregon voters will pass Measure 2. Regardless of how you yourself voted on it, do you think Measure 2 will pass, or will it fail? What percentage of the vote will the YES side get, given that you think Measure 2 will . . . (pass) (fail)?

In all interviews with voters, interviewers were instructed to try to get an exact percentage estimate, rather than a range of percentages.

Table 1. Voters' vs. Politicians' Predictions of Two State Ballot Measures' Outcomes

	Measure 3			*Measure 6*		
	Predict Will Pass	*Predict Won't Pass*		*Predict Will Pass*	*Predict Won't Pass*	
Sample	%	%	*(N)*	%	%	*N*
Voters	36.4	63.6	(275)[a]	38.6	61.4	272[a]
Politicians	26.8	73.2	(56)[a]	10.5	89.5	57[a]
Total	34.7	65.3	(331)	32.6	67.4	(329)
	$\chi^2 = 1.484$, Ns, with 1 df			$\chi^2 = 15.385$, $p < .001$, with 1 df		

	Estimated % "Yes" Votes			
	Measure 3		*Measure 6*	
	%		%	
Actual	49.5		44.9	
Voters' mean estimate	45.8*	$N = 277$	46.7	$N = 271$
Politician's mean estimate	48.1*	$N = 56$	45.2	$N = 56$

	Variance in Estimates	
Voters	162.8686**	155.5001**
Politicians	10.0299**	20.3852**

* Means differ at $p < .05$ ($t = 2.54$, using separate error estimates and reduced degrees of freedom).

** On each measure, voters' variance significantly ($p < .001$) larger than politicians' by F-max test.

[a] Numbers of cases will vary from table to table because of incompletes.

vote, carrying most counties by a narrow margin. The 1984 result was very similar, the tax limitation measure receiving 49.3 percent of the vote and once again carrying most counties in the state. In contrast, the 1982 land-use measure got only 44.9 percent statewide and was defeated in most counties. Unless specifically noted, all comparisons reported here are from the 1982 mail and exit surveys.

Table 1 provides several tests comparing the accuracy of voters and politicians on each ballot measure. Majorities of both groups said each measure would fail, and all their estimated percentages had means of less than 50 percent. Nevertheless, all between-group comparisons favored the politicians: In all four cases, politicians' predictions were closer to the outcome than were voters'. For the property tax measure, politicians were only marginally more likely than voters to think the measure would fail, but their mean percentage estimate was significantly closer to the outcome ($p = .011$). As for the land-use measure, politicians were much more likely than voters to predict its defeat. The mean estimate for politicians was only 0.3 percentage points above the actual outcome (44.9 percent), while voters were 1.8 percentage points too high. Over half the politicians' estimates came within 3 percentage points of the actual result for the land-use measure, compared to less than one in five voters ($p < .001$ by chi-square test).

Variances in voters' percentage estimates were much higher than for estimates by politicians, requiring both *t*-test comparisons in Table 1 to assume heterogeneity of variance. Voter estimates ranged all the way from 2 to 91 percent.

This much greater range among voters' estimates suggests another possible explanation for politicians' greater accuracy: Perhaps accuracy differences reflect politicians' greater knowledge and sophistication, rather than constraint. Most sophisticated observers of U.S. elections know that almost all outcomes will fall between 40 and 60 percent. In this study, zero and 3.6 percent of politicians fell outside that range, compared with more than 30 percent (tax) and more than 25 percent (land use) of the voters.

Judging by a succession of tests, however, it won't be easy to explain voter-politician differences by means of the sophistication argument. First, even if we delete from comparisons everyone whose estimates fell outside the 40–60 percent range, voters' mean estimates hardly changed at all. A second variant of the sophistication argument notes that voters tended to use "round numbers" (i.e., multiples of five) in their estimates, while politicians used numbers in between, thus presumably putting politicians' estimates closer to the real-number outcome of 49 percent for the tax measure. Even after we delete both all cases outside the 40–60 percent range and all estimates made in multiples of five, accuracy differences remained robust. As with all the previous tests of percentage estimates, politicians' estimates clustered much more closely around the actual results.

What about the Looking-Glass effect? Were there any apparent signs of it, and if so, did the effect seem stronger among voters than politicians? The answers seem to be yes—and yes. We asked voters to tell us how they voted after they had made their predictions, but we did not ask politicians directly about their own attitudes toward the two ballot measures. For politicians, the best indicator we have of their own attitudes was their affiliation with either the Democratic or Republican parties. Generally speaking, both Oregon ballot measures were strongly supported by conservative groups whose membership overlapped with major elements of the Oregon Republican Party, and were strongly opposed by almost all Democratic politicians in the state. (Many moderate Republicans also opposed both measures, however.)

Let us consider Looking-Glass results for voters first, for politicians second, and then compare the apparent strength of the effect between the two groups. Table 2 presents all these results.

Among voters, how they voted very strongly predicted a belief that

Table 2. Apparent Looking-Glass Effects Among Voters, Politicians

	Measure 3 (Tax)			*Measure 6 (Land Use)*		
	Predict Will Pass	*Predict Won't Pass*		*Predict Will Pass*	*Predict Won't Pass*	
	%	%	*(N)*	%	%	*(N)*
Subsample						
Voted "yes" on this measure	70	30	(60)	74	26	(82)
Voted "no" on this measure	25	75	(194)	19	81	(189)
Rep. pol.	48	52	(23)	17	83	(24)
Dem. pol.	9	91	(32)	6	94	(32)

χ^2 Comparisons	*Measure 3*	*Measure 6*
"Yes" voters vs. "no" voters	$\chi^2 = 39.07, p < .001$, with 1 df	$\chi^2 = 65.11, p < .001$, with 1 df
Rep. vs. Dem. pol.	$\chi^2 = 8.49, p < .01$, with 1 df	$p < .20$ by Fisher's Exact Probability Test
"Yes" voters vs. Rep. pol.	$\chi^2 = 2.65, p < .15$, with 1 df	$\chi^2 = 23.71, p < .001$, with 1 df
"No" voters vs. Dem. pol.	$\chi^2 = 2.89, p < .10$, with 1 df	$\chi^2 = 2.84, p < .10$, with 1 df

	Mean Estimated Percentage "Yes"			
	Measure 3		*Measure 6*	
Subsample	%		%	
Voted "yes" on measure	51.15^{a}	($N = 61$)	$53.96^{a,c}$	($N = 83$)
Voted "no" on measure	$44.16^{a,b}$	($N = 194$)	43.27^{a}	($N = 167$)
Rep. pol.	49.00^{d}	($N = 24$)	$46.33^{c,e}$	($N = 24$)
Dem. pol.	$47.26^{b,d}$	($N = 31$)	44.26^{e}	($N = 31$)

[a] Means differ at $p < .001$, based on t-test using pooled error estimate.

[b] Means differ at $p < .01$, based on t-test using separate error estimates and reduced degrees of freedom.

[c] Means differ at $p < .001$ based on t-test using separate error estimates and reduced degrees of freedom.

[d] Means differ at $p < .05$, based on t-test using pooled error estimate.

[e] Means differ at $p < .10$, based on t-test using pooled error estimate.

their own side would win, regardless of whether their own vote was "yes" or "no" and regardless of which measure it was. Looking across both measures, at least 70 percent of those who voted "yes" also thought their measure would win, while at least 75 percent of "no" voters thought that measure would be defeated ($p < .001$ for each ballot measure). Similar findings occurred for estimates of percentages. "Yes" voters estimated "yes" votes at a mean 7 percent higher (tax measure) and 10 percent higher (land use) than the means for "no" voters ($p < .001$ in each case).[3]

As for the politicians, the evidence for a Looking-Glass effect was stronger for the tax relief measure. Republicans were much more likely to think that the tax limitation would win ($p < .01$), and their mean percentage estimate on the measure also was higher than the Democrats' ($p < .05$). Results were in the same direction for the land-use measure, but not significantly so.

Already the Looking-Glass evidence among politicians seems weaker than among voters. Is there a tougher and more direct way, though, to compare the two? One highly plausible way would be to compare those whose attitudes ought to be distorting their perceptions in the same directions. If, for example, Democratic politicians are closer to the actual results than are "no" voters, we could argue that Democratic politicians' Looking-Glass effect was the weaker of the two. Then we would do a similar comparison between Republican politicians and "yes" voters. Table 2 again contains the results bearing on these comparisons. Looking first at the question of whether the measure will win or lose, we find that in all four cases ("yes" voters vs. Republican politicians and "no" voters vs. Democratic politicians, on each ballot measure) the politicians less often predicted the measure would win. Looking now at the bottom of Table 2, we find that the politicians are closer to the actual percentage outcome in all four cases. Summing all these comparisons across both percentage estimates and "win-lose" questions, in all eight cases politicians' predictions were closer to the mark, and five of these eight differences reached at least the $p < .10$ level.

A slightly different way of comparing the strengths of apparent Looking-Glass effects would be to look at the *distances* of estimates from actual two-digit percentage outcomes, once again comparing Republican politicians with "yes" voters and Democrats against "no" voters. For each 1982 ballot measure, exactly half the Republican politicians made estimates that were 3 percentage points or less away from the actual results, compared to less than one in four (taxes) and one in nine (land-use planning) "yes" voters. These differences were significant at $p < .05$ and $p < .001$, respectively, by chi-square tests. Nearly three-fourths of the Democratic politicians were within 3 percentage points of the tax outcome, compared to less than 18 percent of the "no" voters ($p < .001$). As for the land-use outcome, more than 61 percent of the Democratic politicians were within 3 percentage points of the outcome, while only 27 percent of the "no" voters were ($p < .001$).

---

[3] The same apparent Looking-Glass effect was present among 1984 voters—and at least as strongly. Some 74 percent of the "yes" voters on the property tax measure thought it would pass, while 85 percent of the "no" voters thought it would be defeated. In addition, "yes" voters estimated "yes" votes at a mean that was 53.8 percent while "no" voters' mean estimate was only 41.2 percent (all differences significant at $p < .001$, by χ^2 and t-tests, respectively).

If politicians were better predictors than the voters, was it because politicians had more public opinion data to aid them? The top half of Table 3 compares the numbers and types of data sources used by politicians (in 1982) and voters (in 1984). In each survey, respondents could cite as many as 10 possible sources of public opinion data. Though the two lists of 10 sources weren't completely identical,[4] comparison of the items they did have in common (e.g., published polls, media reports and letters to the editor) clearly shows that politicians used them much more often than did voters (Table 3). Overall, politicians used a mean of 4.9 of their 10 sources, while voters used only 2.2 of their 10.

Apparently, then, politicians used a greater variety of sources to make their predictions than did voters. It is tempting to test whether the greater variety of sources used by politicians (in 1982) made them better predictors than voters were in 1984. But when the question is put this way, the hazards of assuming complete comparability are obvious. Not only were the lists of sources not identical for the two elections but, in 1982, the sources were for *two* ballot measures while in 1984 they were for only one. Nevertheless, we can stay within each data set and use sources to predict accuracy differences *among* politicians in 1982 and *among* voters in 1984.

When we look separately at each one of the 10 different sources, only one—published polls[5]—predicted accuracy differences among politicians as well as among voters (see bottom half of Table 3). For

---

[4] It was inevitable that the two lists of sources would not be identical because there was no real equivalent, for voters, of such obvious and important sources for politicians as contacts with constituents or impressions about voters from political colleagues. In order to try to be fair to voters, however, we separated such similar sources as "conversations with family and friends" from "conversations with acquaintances," so that one basic type of source—conversations—actually could produce a count of two sources for voters.

[5] The state's most prominent pollster, J. Roy Bardsley, told the author that his final published polls showed all three measures getting a plurality of "no" votes. However, in each case the number of undecided voters was larger than the margin of apparent defeat. Further, the time lags between final polling and election days ranged from one to two weeks. Final Bardsley and Haslacher, Inc. figures for 1982: Measure 3—37 percent yes, 45 percent no, with 18 percent undecided; Measure 6—38 percent yes, 40 percent no, with 22 percent undecided. The firm's final figures for the 1984 tax measure were 46 percent yes, 49 percent no, with 5 percent undecided. According to Bardsley, his published poll was the only one to predict the 1984 defeat of the tax measure.

Table 3. Sources of Data for Predictions

Politicians' Sources 1982	*% Using (N = 58)*	*Voters' Sources 1984*	*% Using (N = 408)*
Personal contacts, conversations	89.7	Family, friends	58.3
Published polls	77.6	Polls in news	15.4
Media reports	69.0	News reports	40.7
Contacts with groups	60.3	Meetings	4.2
Informants, colleagues	58.6	"Experts"	12.5
Editorials	37.9	Editorials and advertisements	30.4
Letters to the editor	29.3	Letters to the editor	12.5
Constituent letters & calls	27.6	Acquaintances	15.9
Private polls	22.4	Bumper stickers, lawn signs	11.8
Other	12.0	Other	13.2
Mean sources named: 4.862		Mean sources named: 2.167	

Sources as Predictors of Accuracy

Will 1982 tax limit pass?	Yes	No	Will 1984 tax limit pass?	Yes	No
Used polls[a]	24.4%	75.6%	Used polls[a]	29.8[b]	70.2[b]
Did not use polls[a]	30.8	69.2	Did not use polls[a]	41.0[b]	59.0[b]
Will measure 6 pass?	Yes	No	% estimate (means)	Tax Limit	
Used polls	4.4%[c]	95.6%[c]	Used polls	49.50%	"Yes"[d]
Did not use	30.8[c]	69.2[c]	Did not use	43.77%	"Yes"[d]
% estimate (means)	Tax Limit	Land Use	Used bumper stickers, lawn signs	40.55%	"Yes"[f]
Used polls	47.95%	44.74%[e]	Did not use	45.21%	"Yes"[f]
Did not use	48.25%	46.67%[e]			

[a] "Polls" always means published polls; 1982 respondents: politicians; in 1984: voters.
[b] By χ^2 test, $p < .10$.
[c] Fisher's Exact Test, $p = .0106$, two-tailed.
[d] Means differ at $p < .01$, $t = 3.05$.
[e] Means differ at $p < .001$, $t = 3.55$.
[f] Means differ at $p < .05$, $t = 2.20$.

politicians, using polls seemed to work only in predicting the land-use measure. With regard to the 1982 tax limitation measure, politicians who used the polls were in fact very slightly farther away than their colleagues from the actual percentage. Interestingly, six politicians carried their tax measure estimates to three decimal places and every one of them very precisely bracketed the actual outcome of 49.5 percent. None of the 10 sources—not public polls, not private polls, not any—differentiated these six politicians from their colleagues.

Turning to the voter sample from 1984, we find that relatively few voters cited published polls. But those who did were somewhat ($p < .10$) more likely to think the tax measure would fail, while at the same time their higher percentage estimate (a mean of 49.5) was significantly closer to the actual outcome (49.3 percent).

Of all sources used by politicians and voters, only one other source besides polls was significantly related to any measure of accuracy for either group. In this single case, though, this source (lawn signs/bumper stickers) was related to *greater error* among voters.

For the voters, relying on bumper stickers/lawn signs as a source might help "cancel" the benefits of following the polls, when both kinds of sources were used. This pattern of counteracting sources might help explain why more sources of information were not necessarily better, in terms of predictive accuracy. The total numbers of sources used did *not* predict accuracy differences among either the voters or the politicians.

Discussion

This study is the first directly comparing the accuracy of election predictions by politicians with those by members of the mass voting public. As predicted, politicians were better than voters at predicting the outcomes of two controversial ballot measures, and they displayed fewer and weaker signs of the Looking-Glass effect. By showing consistently lower variances among politicians than voters in their percentage estimates, the findings also suggest that it may be more important, first, to consider the pressures—or attributes—politicians have in common, instead of first trying to differentiate among them. The findings suggest also that it is premature to discard models of representative democracy that presuppose politicians' ability to make accurate judgments about constituent opinion.

To show that political decision makers' perceptions of public opinion are more accurate is to support Lemert's argument that decision makers' perceptions are in enough touch with political "reality" that their perceptions will change as the information coming to them about citizen attitudes changes. This proposition thus allows Lemert to argue that the mass media more often change "public opinion" through changes in the information reaching decision makers about citizen attitudes than they do through attitude change itself (see especially Lemert, 1981: 32–37, 213–217).

Why would politicians be less susceptible than voters to wishful thinking? Perhaps, instead of being constrained to avoid wishful thinking, politicians are simply more interested in, and savvy about, elections. Another way of putting this is that we might be comparing full-time professionals against amateurs in the prediction game. Interest alone does *not* seem to provide the explanation. Politicians' personal greater interest predicts their using a greater variety of indicators than voters for their predictions. This appears to have happened. But the total number of different indicators used bore *no* relationship to predictive accuracy within either the voter or the politician data sets. It may be that only some of these indicators enhance predictive accuracy, while others interfere with it.

It was true that using published preelection polls predicted greater accuracy, both among voters and among politicians. And it also was true that politicians were more likely than voters to use published polls. But if some other indicators can interfere with accuracy (bumper stickers/lawn signs for voters), it is hard to see how greater interest alone could lead *both* to using polls and *ignoring* certain of these other indicators. Perhaps interest needs to be accompanied by experience and savvy, both of which could lead to a more judicious use and interpretation of the available evidence.

We can't resolve here whether constraint or savvy—or both—explains why politicians were better able to resist wishful thinking. One reason we can't settle this issue is that greater constraint implies greater pressure toward accuracy, which in turn implies both interest in, and the seeking out of, preelection data. Further, a more judicious, savvy use of these data may itself be a constraint or "survival" characteristic for a politician (e.g., Converse et al., 1966:24; Hedlund and Friesema, 1972:406–7). One of many ways to measure constraint as an independent variable would be to contrast the accuracy of politicians' predictions by the salience of issues: The more salient the issue, the more accurate should be the prediction by politicians. In contrast, the more salient the issue to voters, the more likely that the wishful thinking/Looking-Glass effect might be.

Another issue to be considered here is whether the apparently weaker Looking-Glass effect among politicians could be an artifact of a less precise measure of politicians' attitudes toward the two measures. Suppose that votes cast for and against each measure reflected voter attitudes very precisely, but party affiliation much less precisely categorized politicians' own attitudes. Presumably, then, the relationship between own attitudes and predictions would appear to be weaker among politicians than among voters. It was indeed true that some moderate Republican politicians did oppose one or both of the two measures, and we had no way of identifying them because the returned questionnaires were anonymous. Therefore, these moderate Republicans were combined with more conservative Republicans, most of whom *did* support both measures, thus weakening any signs of wishful thinking among Republican politicians. While plausible, this argument simply cannot explain why Democrats—almost universally opposed to both measures—*also* showed a smaller Looking-Glass effect than did their comparison group of "no" voters.

Finally, let us consider one last threat to these results. The property tax measure would have been adopted as part of the Oregon Constitution had Eugene-Springfield voters not delivered a resounding margin of "no" votes that just barely offset the rest of the state. Could the fact that most of our sample of Eugene-Springfield voters were opposed to both measures have somehow produced the voter errors reported here? No, not at all. In fact, since both measures lost, whatever "bias" existed in our sample usually stacked the deck in the direction of *minimizing* voter error, not enlarging it. If we had weighted our sample of "yes" voters to reflect the overall state outcomes on each measure, we would have greatly *enlarged* voters' predictive errors in three of the four possible cases. Of the only six people to carry out their 1982 estimates to three decimals, every one of them did it only for the tax measure—and every one of them hit the actual outcome, almost to the fourth digit. That all six of them were politicians cannot be explained by either weighting or not weighting a sample of voters.

References

Achen, Christopher H.
  1977 "Measuring representation: perils of the correlation coefficient." American Journal of Political Science 21: 805–15.
Converse, Philip E., Aage R. Clausen, and Warren E. Miller
  1966 "Electoral myth and reality: the 1964 election." In Edward C. Dreyer and Walter A. Rosenbaum (ed.s), Political Opinion and Electoral Behavior: Essays and Studies. Belmont, CA: Wadsworth.
Fields, James M., and Howard Schuman
  1976 "Public beliefs about beliefs of the public." Public Opinion Quarterly 40: 427–49.
Hedlund, Ronald D., and H. Paul Friesema
  1972 "Representatives' perceptions of constituency opinion." Journal of Politics 34: 730–52.
Lemert, James B.
  1981 Does Mass Communication Change Public Opinion After all? A New Approach to Effects Analysis. Chicago: Nelson-Hall.
Miller, Warren E., and Donald E. Stokes
  1963 "Constituency Influence in Congress." American Political Science Review 67: 45–56.
Noelle-Neumann, Elisabeth
  1974 "The spiral of silence: a theory of public opinion." Journal of Communication 24: 43–51.
  1984 The Spiral of Silence: Public Opinion—Our Social Skin. Chicago: The University of Chicago Press.
O'Gorman, Hubert, and Stephen L. Garry
  1976 "Pluralistic ignorance—a replication and extension." Public Opinion Quarterly 40: 449–58.

Election Predictions: An Empirical Assessment

WILLIAM BUCHANAN

POLLSTERS have been publishing predictions of national election results for nearly 50 years. There are now enough instances to make possible a comparison of the *empirical* with the *theoretical* standards of accuracy.

Accuracy of election predictions may be expressed by several measures, each useful for a different purpose. These were developed by the Social Science Research Council team which appraised the performance of the American pollsters after their 1948 fiasco (Mosteller, 1949: 54–61). Among these measures are the following:

1. The proportion of polls which correctly predicted the winner.
2. "Average error," the mean of the differences between the predicted percentage and the actual vote percentage, ignoring the sign. Most easily understood by newspaper readers, this average has the defect of concealing biases—all the prediction errors might be in favor of the more conservative party, or of the ultimate winner or of some other category.
3. Systematic error, the mean of the differences between the predictions and the election percentage, taking account of sign. This reveals the biases, but understates the size of the errors, since large overestimates cancel out equally large underestimates.
4. Nonsystematic error, the standard deviation of the individual er-

---

Abstract Examination of 155 poll forecasts in 68 national elections since 1949 shows that errors average nearly twice what statistical theory would indicate. Polls predict the division of vote between major parties better than individual party percentages, leading to 85 percent success in picking the winner. The worst failures occurred in a few elections where most polls went wrong. Liberal party votes are correctly forecast, conservatives slightly underestimated. Improved polling methods have not led to better forecasts.

William Buchanan is Professor of Politics at Washington and Lee University.

Public Opinion Quarterly Vol. 50:222–227 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-222/$2.50

rors around the mean of the errors in a series of elections. It tells how widely the predictions vary, discounting systematic error. It is not too meaningful to lay readers, but useful to statisticians since it permits one to compare the performance of polls with the deviations that would be expected for given sample sizes. It is these probability calculations that underlie the familiar "margin of error" statements that appear in all responsible reportage of poll forecasts. However, these calculations are based on probability theory, and do not take account of the actual performance of the polls.

These error calculations may be made for each party's vote as a percentage of the total vote, or as a percentage of the "major party" vote. The latter is what most consumers of polls are interested in, since it has political relevance. It predicts who will come in first and how far ahead of the second-place party.

The SSRC's report showed that through 1948 the polls' nonsystematic error was not excessive, but there had been a decided pro-Republican bias, as reflected by the systematic error measure.

Robert M. Worcester (1983) has recently edited a history of political polling that brings together actual and predicted results of 96 election polls in Australia, France, Germany, Britain, Japan, and the Netherlands. I have updated and supplemented these with Canadian, New Zealand and Australian, and U.S. predictions, bringing to 155 the total of forecasts made in 68 elections since 1949. These were the final published percentage predictions of the popular vote for a candidate, party, or coalition in national elections. The data set includes a large majority of such predictions made during the period. Sample size is not always given, but where it is, 1500 or 2000 is the norm, with occasionally larger and rarely smaller samples, none under 900. The assumption for purposes of analysis that the average sample size is 1500 is conservative.

Using the categories established in the SSRC report, average, systematic, and nonsystematic errors were calculated for the party that won the election and also for the one that came in second, each as a percentage of the total vote. Then the percentages of winner and runner-up were added and the winner's percentage of the two-party vote compared to the prediction. This leaves out a residue of votes for minor parties, plus, in a few instances where it was so published, an "undecided" percentage, these being of little interest to consumers of predictions. Since there were several competing forecasts in most elections it seemed sensible to subtract the predicted percentage from the actual percentage; thus errors with a *positive* sign in the calculations that follow are *under*estimates and those with a minus are overestimates of what that party actually got.

The published "margin of error" percentages are derived from the statistical formula for the standard error:

$$1.96\left(\sqrt{\frac{pq}{n}}\right)$$

where p is the proportion we are trying to estimate, $q = 1 - p$, and n is the number of cases in the sample.

The statistical analog for this calculation is a vat containing millions of marbles (equivalent to the electorate) of which about half are white (the leading party's vote). The pollster scoops up a bucketful of marbles (roughly 1500) time after time, counts them, throws them back, stirs the vat and scoops up another bucketful (for 155 samples). If done with marbles we could count on close to 95 percent of the sample estimates varying from the proportion of white marbles in the vat by no more than than the margin of error based on the formula above, in this case 2.5 percent: $1.96\sqrt{(.5 \times .5) / 1500} = .025$.

Merely describing this ideal sampling arrangement is enough to show the contrast to what election forecasters are really doing. Voters aren't susceptible to being stirred and replaced; similar ones are clumped together and surveys sample these neighborhood or telephone prefix clusters. Some aren't home and some don't have phones. Some haven't made up their minds and some won't vote, though they may think they will (we don't know what color marbles these are, or indeed whether they are marbles at all). Each poll has its own adjustment procedure, often a valued trade secret, for coping with these problems. The customary "margin-of-error" stipulations in the media are based upon the formula for simple random sampling, which rounds off to plus-or-minus 2.5 percent for 1500 cases, 2 percent for 2000 cases. With adjustment for cluster sampling these figures would be around 3.5 percent and 3 percent, respectively.

The 95 percent probability criterion calls for an *average* of one "bad" prediction—i.e., one outside the specified margin—in every 20 attempts. National elections are infrequent and we need the full sample of 155 to make use of the principle, which would call about 8 errors of this size.

The calculations of the relevant averages appear in Table 1. The average absolute error for the 155 predictions is close to 2 percentage points—not a bad record. For winning parties the systematic error is +0.45. Thus the winner's percentages were *under*estimated by half a point on the average. Fortunately, the polls also underestimated the percentage of the second party by 0.29, which gave them a better record for their two-party predictions of −0.04—a trivial overestimate of the winner's percentage.

Table 1. Accuracy in Percentage Points of 155 National Poll Predictions, 1950–1984.[a] (Negatives are overestimates.)

	Winning Party	*Runner-up*	*Two-Party Split*
Average error	2.02	2.14	2.16
Systematic error	+0.45	+0.29	−0.04
Nonsystematic error	2.62	2.85	2.73
	Conservative		*Liberal/Labor*
Systematic error	+0.89		−0.08
	Systematic	*Nonsystematic*	*N*
By decade			
1950s	0.09	2.89	13
1960s	−0.39	2.15	29
1970s	0.15	2.60	67
1980s	−0.12	3.14	46

[a] Australia 20, Canada 16, France 26, West Germany 9, Japan 14, Netherlands 1, New Zealand 4, United Kingdom 36, United States 29.

The nonsystematic error is an indication of how closely these predictions cluster around the actual percentages, *after* allowing for these biases. This error component may be thought of as the net effect of all those uncorrected-for not-homes, nonvoters, undecideds, and mind-changers, plus random sampling errors, with some of the errors canceling out others. We may compare the standard deviations for these elections with the theoretical ones derived by the statisticians' formula, which represents the error due to sample size alone. Standard error from the formula above is $\sqrt{pq/n}$. The vote for the average winning party is 47 percent of the total, so we insert .47 for p and 1500 for n, our approximation of the average sample size, to get a theoretical standard error of 1.29 percentage points, which should be comparable to our nonsystematic errors.

But it is not. Standard deviation of the errors for winning parties turns out to be 2.62 percentage points, twice what would be expected. Multiplying by 1.96 we get a margin-of-error estimate of 5.1 percent, which we may compare to the *theoretical* margins of 2.5 percent to 3.5 percent as calculated above. Where theory would lead us to expect 7 or 8 errors larger than 3.5 percent, there were in fact 31 errors of this magnitude in the two-party predictions. As Paul Perry (1979) of the Gallup organization has observed: "One cannot expect to achieve in surveys an accuracy consistent with the theory of random sampling." On the basis of the long-term performance of the polling profession as a whole, press accounts of predictions should say "5 percent margin of error" where they now say "3 percent" and make comparable adjustments to all such estimates based on sample size.

Ideologues of left and right are convinced that pollsters rig their predictions to favor the other party. There is a greater likelihood, however, that the method itself is biased toward one side or the other. Perhaps the most perplexing problem faced in the adjustment decision is how to screen out those citizens who express an intention but are unlikely to reach the voting booth to implement it. Pollsters do this by eliminating or down-weighting on the basis of other questions about past voting habits, registration, interest in the election, and so on (Lipset, 1980; Traugott and Tucker, 1984; Wiener, 1976). If the screen is too loose, the intentions of nonvoters will be counted; if it is too tight some potential voters will be left out. There are also socioeconomic biases in sampling: e.g., the rich can restrict interviewers' access to their dwellings and can have unlisted phones; the poor are suspicious of interviewers and also may have no phones. The problem of election prediction is whether these screening, sampling, and other biases cancel out or whether they cumulate, as they obviously did in 1948.

We can shed some light on the question by grouping the parties into the more conservative ones (winners 103 times) and liberal or labor parties (winners 52 times) and comparing the systematic errors. It appears that there was an underestimate of conservative votes by nearly a percentge point (+0.89) and a trivial overestimate of the liberal/labor vote (−0.08). These led to predicting a liberal victory 17 times when the conservatives won (17 percent wrong) compared with 5 bad forecasts when the liberal or labor party won (10 percent wrong).

We might expect the predictions to have improved over the years since 1949, particularly in the 1980s, when random-digit-dialing and widespread telephone ownership have made possible forecasts based on data collected only a day or two before the election. Apparently not. The results by decades in Table 1 show no consistent pattern by either error measure.

In general the most serious errors occurred in a few elections in English-speaking democracies where several polls made the same mistakes. In the British election of 1970, five polls overestimated the Labor vote by an average of more than 4 percentage points, leading to a two-party error of some 3.5 points in favor of a party that lost the popular vote by 2.5 points. Thus 4 out of 5 forecasts picked the wrong winner. The year 1983 was in some respects even worse; the average of six polls was an overprediction of the Conservative vote by over 3 points and an *under*prediction of Labor by about 2 points, resulting in a two-party bias toward the Tories of nearly 4 points. But Mrs. Thatcher had such a commanding lead that no prediction was wrong and the error attracted less comment than the smaller one in 1970. Bad years in Australia were 1980, when a 3-point average underestimate of the Liberals caused five polls to go astray in a very close election, and 1984, when three out of four polls predicted a much wider margin for Labor than eventuated. In 1976 and 1980 in the U.S., both major parties were seriously underestimated, but the underestimates offset each other in the two-party percentages and put most of the polls on the right side. In all, 22 predictions (out of 155) pointed to the wrong winner, half of them occurring in these debacles.

The consistency with which almost all the polls in a few elections made errors in the same direction and of about the same magnitude is revealing. These are situations in which the electorate is volatile and people make up or change their minds in the last few days about whether and/or how they will vote. Survey procedures based on experience in more stable situations may not work as well. The popular stereotype that "the polls are always wrong" is incorrect. They are on the wrong side in only one prediction out of every seven. A better oversimplification would be: "Whenever the polls are wrong, they are all wrong."

To summarize, examining the performance of most of the published national election predictions since 1949 in a variety of political cultures and electoral systems shows: (1) The margin of error estimates based on probability theory seriously understate the empirical range of error. (2) Winners' percentages are underestimated on the average by about half a point and runners-up by a quarter point. (3) These compensating underestimates lead to better predictions of the two-party distribution and hence to picking the winner about 85 percent of the time. (4) There is no evidence that either biases or random errors have been reduced over the years. (5) There is a tendency to underestimate the conservative vote. (6) The bad predictions are made in a few elections when most of the polls err in the same direction.

References

Lipset, Seymour Martin

1980 "Different polls, different results in 1980 politics." Public Opinion, August/September, p. 19.

Mosteller, Frederick, Herbert Hyman, Phillip J. McCarthy, Eli S. Marks, and David B. Truman

1949 The Pre-election Polls of 1948. New York: Social Science Research Council.

Perry, Paul

1979 "Certain problems in election survey methodology." Public Opinion Quarterly 43:322–25.

Traugott, Michael W., and Clyde Tucker

1984 "Strategies for predicting whether a citizen will vote and estimation of political outcomes." Public Opinion Quarterly 48:330–43.

Weiner, Sanford L.

1976 "The competition for certainty: the polls and press in Britain." Political Science Quarterly 91:673–696.

Worcester, Robert M.

1983 Political Opinion Polling. New York: St. Martin's Press.

The Presence of Others and Overreporting of Voting in American National Elections

BRIAN D. SILVER, PAUL R. ABRAMSON
AND BARBARA A. ANDERSON

SURVEY RESPONDENTS often exaggerate their conformity with socially approved norms. To improve the quality of survey data, it is important to establish under what conditions people are more or less likely to exaggerate such conformity. Unfortunately, there are very few external standards against which reported behavior can be validated. Voting differs from most forms of political behavior because self-reports of voting can be compared with official voting statistics. Postelection sur-

---

Abstract This study uses the 1978 and 1980 vote validation studies conducted by the University of Michigan Survey Research Center to test the extent to which false claims about voting are affected by the presence of third parties during the interview. The presence of third parties during interviews is far more frequent than is commonly assumed. But the tendency of respondents to give socially approved answers is not affected by the presence of others during the interview. Thus, additional efforts to avoid contamination of interviews by eliminating third parties are not likely to reduce the exaggeration of self-reported vote. The analysis suggests that the declared intention to vote is a far more important factor in whether people falsely report voting than is the presence of others. Additional effort to understand the motivational basis of voting and nonvoting could help to account for variation in voting overreports.

Brian D. Silver and Paul R. Abramson are Professors of Political Science at Michigan State University, and Barbara A. Anderson is Professor of Sociology and Research Associate of the Population Studies Center at the University of Michigan. The data for this research were made available by the Inter-University Consortium for Political and Social Research, which bears no responsibility for the authors' analyses or interpretations. The authors are grateful to Mary R. Jackman, Robert W. Jackman, Howard Schuman, and Arland D. Thornton for helpful comments on an earlier draft; to Harriet Dhanak for programming assistance; and to Maria Elena Sanchez for providing information concerning the interviewer instructions in SRC National Election Studies. Work on this topic was motivated by research on the Soviet Interview Project, which is funded by Contract #701 from the National Council for Soviet and East European Research to the University of Illinois (James R. Millar, Project Director).

Public Opinion Quarterly Vol. 50:228–239 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-228/$2.50

veys in the United States consistently show a substantially higher proportion of the population claiming to have voted than the official vote records reveal.

In an attempt to determine the extent and patterns of this bias, the University of Michigan Survey Research Center conducted validation studies in 1964, 1976, 1978, and 1980, which allow comparison between self-reported voting and official voting records of the survey respondents. In these four studies, between 20 and 30 percent of the respondents who did not vote, according to the validation checks, claimed to have voted.

Although there have been several published articles devoted to the study of vote validation (e.g., Clausen, 1968–69; Katosh and Traugott, 1981; Sigelman, 1982), they have not shed much light on the conditions that contribute to false reporting of voting. Moreover, if respondents overreport their voting, they may overstate their participation in other socially desirable political activities, such as attending political meetings and contributing money to candidates. Although the actual extent of exaggeration by individuals of behaviors other than voting cannot be easily directly determined by comparisons with external criteria, understanding the sources of error in voting reports might lead to improved reliability of other reports as well.

Sudman and Bradburn (1974:1–27) and Bradburn (1983) identify three principal sources of variation in survey data quality: (1) the characteristics of the task (the survey instrument and context), (2) the interviewer's performance, and (3) the respondent's characteristics. Studies of vote validation have focused almost exclusively on the third source: respondent characteristics.[1] These studies have shown that overreporting of voting occurs in all social categories of the population, and they have concluded that the probability of overreporting is very weakly related to such respondent characteristics as socioeconomic status, age, and sex, or to such political attitudes as party identification, the sense of political efficacy, feelings of political trust, or sense of citizen duty.

The only social background characteristic that is reported to be strongly related to the size of the discrepancy between self-reported and validated vote is race: Blacks appear to be consistently more likely to overstate voting than whites (Traugott and Katosh, 1979; Katosh and Traugott, 1981; Sigelman, 1982; Abramson et al., 1983; Abramson and Claggett, 1984; Hill and Hurley, 1984).[2] However, given the far larger number of whites than of blacks, four-fifths of the respondents who overreport voting are whites.

---

[1] One exception is the research based on the Denver Validity Study of the late 1940s, which also examined interviewer effects. See Feldman et al., 1951; and Cahalan, 1968–69. Another is Weiss's (1968–69) analysis of the validity of voting reports (and other reports) of a sample of New York welfare mothers.

The main implication of previous research is that overreporting of voting is almost entirely a result of interviewer performance, characteristics of the survey task, or a combination of the two. Overreporting rarely results from a failure of memory, since the postelection surveys in which respondents are asked whether they voted in the recent presidential or midterm election are fielded within a few months of the election date, with a majority of the respondents interviewed within two months of the election. Instead, the response error appears to be more of the type obtained in market surveys in which people greatly exaggerate how frequently they brush their teeth (cf. Menneer, 1979:144).

There is some question, however, whether altering the characteristics of the survey task can reduce the tendency to exaggerate voting. There is very little published research on this topic. Research by Bradburn and Sudman (1980:1–13) suggests that altering the *form* of the survey, such as by using self-administered or telephone surveys or random response techniques, only slightly reduces the tendency of respondents to exaggerate their performance of socially desirable acts.

Other features of the survey task might contribute to the tendency toward overreporting of voting. If overreporting of voting is caused by the tendency of respondents to exaggerate their performance of socially desirable behavior, then the tendency toward exaggeration might be affected by whether people are present during the interview. The effect of the presence of third parties during interviews has rarely been studied.

Bradburn and Sudman report one study that compared the patterns of responses to attitudinal items between people who were interviewed in isolation and people who were interviewed in the presence of third parties. They found no consistent effect of third persons on the response patterns (Bradburn and Sudman, 1980:134–46). But Bradburn and Sudman's study did not test for the effects of third parties on responses whose accuracy could be *externally* validated. The SRC vote validation studies provide the opportunity for such a test.

---

[2] The Survey Research Center's validation of actual voting is not necessarily error-free. It is plausible that at least part of the discrepancy between reported and validated vote among blacks is due to errors in the records in local voting precincts, i.e., to the omission of some black voters. This seems especially likely to occur in rural areas and in the South. Southern *whites* also show a somewhat greater discrepancy between the number of self-reported and validated voters than do whites from other regions of the nation. See Abramson and Claggett, 1984.

How Is the Presence of Third Parties Likely to Affect Vote Overreporting?

Bradburn and Sudman (1980:134) state, "It is generally considered good field practice not to have anyone other than the interviewer and the respondent present during a face-to-face interview. . . ." One reason for trying to avoid having third parties present during interviews is to minimize contamination or distortion of the respondent's answers.

In the case of reported voting, the likely direction of bias caused by the presence of third parties is uncertain, because it may vary according to whether or not the third party knows whether or not the respondent did in fact vote. If the third party does not know whether or not the repondent voted, the respondent may be more likely to report socially desirable behavior and to claim to have voted even if he or she did not go to the polls. But in many cases the third party may know that the respondent did not vote, making it less likely for a nonvoter to claim falsely to have voted. In addition, parents may be especially reluctant to lie in front of their children.

The extent and direction of the effects of third-party presence is an empirical question, the answer to which rests on two behavioral conditions: (1) the frequency with which third parties are present during interviews, and (2) the impact of the presence of third parties on responses.

How Frequently Are Third Parties Present?

There is little published information about how often the desired goal of isolating the respondent one-on-one with the interviewer is met in practice. Bradburn and Sudman (1980) describe one national survey conducted in 1975 by the National Opinion Research Center (NORC) in which experienced interviewers were given special instructions to try to get all third parties out of the room before interviewing. In that survey, interviews were conducted with third parties present in 26 percent of the cases.

This suggests that when no special effort is made to reduce the presence of third parties, they are even more likely to be present. In Verba and Nie's (1972) *Participation in America* survey, conducted by NORC, third parties age 15 or older were present in 43 percent of the interviews.[3] In Jackman and Jackman's (1983) *Class Awareness in America* study, conducted by the SRC, third parties age 10 or older were present in 42 percent of the interviews.[4] In the Soviet Interview Project General Survey, conducted by NORC, interviews were conducted with third parties present in 52 percent of the cases (Anderson and Silver, 1986).

---

[3] This figure is calculated from the ICPSR codebook for the Verba-Nie study.

In surveys of women in less developed countries conducted for the World Fertility Survey, strict instructions were given to interviewers to administer the interviews privately (Casterline and Chidambaram, 1984:268). Nonetheless, on average in about half of the interviews third parties are reported to have been present.[5]

Unfortunately, few American survey research organizations routinely report information about the presence of third parties during the interviews. For example, NORC's General Social Survey does not gather this information. Fortunately, since 1966 the SRC American National Election Studies have collected information about the presence of others during interviews. This allows us to examine the relationship between overreporting of voting and the presence of third parties for both the 1978 and 1980 SRC Election Studies.[6]

Examination of the marginal distributions for those surveys reveals a very high incidence of third-party presence. As shown in Table 1, in the 11 SRC National Election surveys in which data on who was present during the interviews were collected, the lowest incidence of third parties being present was 40.6 percent (the 1980 postelection survey) and the highest was 56.8 percent (the 1970 postelection survey).

Despite the goal of isolating respondents one-on-one with the interviewers, third parties were present in the SRC National Election surveys during about *half* (51.2 percent) of the interviews. Consequently, if the presence of third parties during interviews affects the tendency to overreport socially desirable behavior, then the *potential* for an effect is considerably higher than is generally assumed.[7]

The effect of the presence of others might depend on the relationship

---

[4] Personal communication from Mary R. Jackman.

[5] On average, in World Fertility survey studies conducted in 24 less developed countries, third parties were present in 47.8 percent of the interviews (calculated from Table 1 of Casterline and Chidambaram, 1984).

[6] The 1964 Election Study conducted a vote validation but did not collect information about the presence of others in the interview. The 1976 Election Study collected information about the presence of others in the preelection survey but not in the postelection survey, when respondents were asked to report whether they had voted. Our treatment of the vote validation data from the 1978 and 1980 Election Studies follows the procedures used by Abramson and Claggett (1984).

[7] The bias would likely depend on whether third parties were present during the portions of the interviews when questions about socially desirable behavior were asked. The SRC National Election studies do not record when the third parties were present during the interviews. The instructions to interviewers state that third parties are to be counted as present if they were present during any part of the interview session.

Table 1. Presence of Others During the Interviews in SRC American National Election Studies, 1966–1982[a]

Who Was Present Other Than Respondent and Interviewer?	*1966 Post*	*1968 Pre*	*1968 Post*	*1970 Post*	*1972 Pre*	*1974 Post*	*1976 Pre*	*1978 Post*	*1980 Pre*	*1980 Post*	*1982 Post*
Nobody	44.3%	45.1%	49.3%	43.2%	44.8%	48.8%	54.3%	48.8%	49.7%	59.4%	55.3%
Children under 6 only	8.4	6.3	7.9	8.0	7.2	8.0	7.1	7.9	5.6	7.2	5.4
Older children or older & younger only	6.0	7.4	7.0	7.9	6.2	6.6	5.7	6.2	5.8	5.2	5.2
Spouse[b]	22.2	21.8	18.4	22.1	20.1	16.3	15.7	18.3	18.1	13.6	16.4
Spouse with children[c]	9.9	9.8	9.5	10.3	10.8	9.3	6.9	8.1	8.7	6.0	8.5
Other adults with or without children[d]	9.1	9.7	7.9	8.5	10.8	11.0	10.3	10.8	12.1	8.6	9.2
Total with somebody	55.7%	54.9%	50.7%	56.8%	55.2%	51.2%	45.7%	51.2%	50.3%	40.6%	44.7%
Base *N*[e]	(1279)	(1535)	(1334)	(1490)	(2690)	(2509)	(2847)	(2290)	(1613)	(1405)	(1416)

[a] Source: Compiled from frequency distributions in codebooks distributed by the Inter-University Consortium for Political and Social Research (Ann Arbor, MI). The table includes all SRC American National Election studies for which data on the presence of others were collected. The presidential year studies included both pre- and postelection interviews; the midterm election studies conducted postelection interviews only. However, information about the presence of third parties was not always collected in both the pre- and postelection surveys for a given year.

[b] Includes cases where other adults (relative or nonrelatives) were also present with spouse.

[c] Includes cases where other adults were also present with spouse.

[d] Includes cases where other relatives or other adults were present, with or without children, but without spouse being present.

[e] The Base *N* excludes respondents for whom the information was not ascertained. Figures for 1974 and 1976 are based on weighted *N*s.

of the third party to the respondent. As shown in Table 1, the most frequent "other person" in the interviews is the respondent's spouse; the second most frequent is a child. Table 2 shows that the probability that a third party will be present during the interview is about the same for blacks as for whites, and the same for males as for females. However, there are differences in the type of other person who is likely to be present.[8] The spouse is much more likely to be present for whites and for males than for blacks and for females. Children are more likely to be present for blacks and for females than for whites and for males.

Presence of Third Parties and Overreporting of Vote

Table 3 shows the percentage of validated nonvoters in 1978 and 1980 who say that they voted, controlling for the presence of other

---

[8] The SRC National Election surveys vary in size from year to year, partly because of the availability of funding and partly because of varying research goals. The 1978 midterm survey (which, like all but one midterm survey, was based only on a postelection interview) was the largest of the eight midterm surveys, mainly because of increased funding to study congressional elections. The 1980 survey (with a total *N* of 1614) was slightly smaller than the average for presidential election surveys, and our *N* is reduced further because the 206 respondents who were not included in the postelection interview are not included in our analysis.

Table 2. Presence of Others During Interviews, by Race and Sex, 1978 and 1980 SRC Studies[a]

	1978				*1980*			
	Race		*Sex*		*Race*		*Sex*	
Other Person Present?	*White*	*Black*	*Male*	*Female*	*White*	*Black*	*Male*	*Female*
Nobody	49.4%	46.8%	47.8%	49.8%	60.4%	53.0%	55.6%	62.3%
Child under 6	7.4	12.0	2.2	12.4	6.9	8.4	2.6	10.7
Older child	6.0	7.7	3.3	8.5	4.7	8.4	2.5	7.3
Spouse	19.4	8.6	25.0	12.8	14.3	8.4	20.9	7.9
Spouse & child	8.1	7.3	11.7	5.3	6.3	4.8	9.2	3.7
Other	9.7	17.6	10.0	11.2	7.3	16.9	9.3	8.1
Somebody (total)	50.6%	53.2%	52.2%	50.2%	39.6%	47.0%	44.4%	37.7%
Base *N*	(2003)	(233)	(1000)	(1263)	(1225)	(166)	(612)	(793)

[a] Source: Figures in this table and subsequent tables are calculated from data in the 1978 SRC American National Election Study and the 1980 American National Election Study postelection survey.

parties during the interview.[9] Nonvoters are slightly more likely to say that they voted when *nobody* other than the respondent and interviewer is present than when somebody else is present, but the difference is not statistically significant.[10] More important, the relationship between vote overreport and *who* was present shows little consistency between the 1978 and 1980 postelection surveys. For example, in the 1978 survey, validated nonvoters were *less* likely to say they voted when their spouse was present during the interview than when their spouse was not present; in the 1980 survey, validated nonvoters were *more* likely to say they voted when their spouse was present.

In both surveys, validated nonvoters were more likely to say they voted in the presence of older children than in the presence of younger children. It is possible that the presence of older children stimulated the respondents to provide a normatively correct response. Even if this is a systematic source of error, however, its aggregate impact is small, since, as Table 1 shows, older children are present in a very small proportion of the interviews. Moreover, the proportion of validated

---

[9] The *N*s for 1978 are substantially greater than those for 1980 for two reasons. First, as mentioned above, the 1978 sample is larger. Second, we are presenting the percentage who claimed to have voted only among validated *nonvoters*, and nonvoting is substantially higher in midterm elections than in presidential elections.

[10] A difference of proportions test (a two-tailed test) between whether Nobody was present and Somebody was present, based on the distributions reported in Table 3, revealed that the *p*-value for 1978 was .646 and the *p*-value for 1980 was .342.

Table 3. Percentage of Validated Nonvoters Who Say That They Voted, by Presence of Others During Interview, 1978 and 1980 SRC Postelection Surveys

		Other Persons Present							
Year		*Nobody*	*Somebody*	*Child Under 6*	*Older Child*	*Spouse*	*Spouse & Child*	*Other*	*All Nonvoters*
1978	%	23.1%	22.0%	22.9%	28.9%	19.5%	17.9%	23.9%	22.5%
	Base *N*[a]	(581)	(673)	(131)	(76)	(205)	(106)	(155)	(1254)
1980	%	29.2%	25.4%	15.5%	24.0%	34.4%	16.7%	31.5%	27.5%
	Base *N*[a]	(267)	(228)	(58)	(25)	(61)	(30)	(54)	(495)

[a] The Base *N* is the total number of validated nonvoters in each subset. The percentage shows the proportion of validated nonvoters who said that they voted.

nonvoters who report voting when children (young or old) are present is not consistently higher or lower than the average extent of overreporting by nonvoters.

There may be a consistent tendency to misreport less when *both* the spouse and child are present. This suggests that there may be a norm against lying in front of children. But the total number of interviews conducted with both the child and spouse present is small, and even the 11 percentage point difference in 1980 between this subset and all other validated nonvoters is not statistically significant. On balance, it appears safe to conclude that reports of voting by validated nonvoters are not affected very much by the presence of others during the interview.

Other Possible Factors Affecting Vote Overreporting

The lack of a consistent relationship between the kinds of people present and the extent of overstatement of voting is strong indirect testimony to the saliency of the civic norm of voting. Many Americans apparently find it difficult to acknowledge that they do not vote. There is at least indirect evidence in the SRC Election Study data that this reluctance to acknowledge not voting derives from the widespread norm that people should vote. One very strong predictor of whether voters were likely to report falsely that they voted—a predictor that has been overlooked in the research literature—is whether respondents stated in the preelection survey that they expected to vote.

Table 4 shows the percentage of validated nonvoters in 1980 who said that they voted, controlling for both the presence of others during the postelection interview and whether or not respondents said they expected to vote during the preelection interview. It shows a striking difference in the extent of vote overreporting between those who expected to vote and those who did not. This difference is far greater than

Table 4. Percentage of Validated Nonvoters Who Say That They Voted, by Whether Respondent Planned to Vote and by Presence of Others During Interview, 1980 SRC Post-Election Survey

		Other Persons Present							
Planned to Vote?[a]		*Nobody*	*Somebody*	*Child Under 6*	*Older Child*	*Spouse*	*Spouse & Child*	*Other*	*All Nonvoters*
Yes	%	53.7%	48.2%	31.8%	46.2%	59.4%	38.5%	53.1%	51.2%
	Base *N*[b]	(136)	(112)	(22)	(13)	(32)	(13)	(32)	(248)
No	%	1.0%	2.0%	0.0%	0.0%	7.4%	0.0%	0.0%	1.5%
	Base *N*[b]	(99)	(99)	(33)	(10)	(27)	(14)	(15)	(198)

[a] The question was worded as follows: "So far as you know how, do you expect to vote in November or not?" Cases where the response was coded Don't Know or Not Ascertained are excluded from the table.

[b] The Base *N* is the total number of validated nonvoters in each subset. The percentage shows the proportion of validated nonvoters who said that they voted.

the effects of the presence of third parties.[11] Respondents who said in the 1980 preelection interview that they planned to vote were much more likely to vote than those who said that they did not plan to vote. But among those who said they planned to vote and did not (*N* = 248), 51.2 percent said that they voted when interviewed after the election. Among respondents who said in the preelection survey that they did not plan to vote and who did not vote (*N* = 198), only three respondents said that they voted.

The consistency between preelection statements of *intention* to vote and postelection *reports* of voting supports the conclusion that variation in vote overstatement is caused in part by variation in the population in the commitment to the civic norm of voting.[12] Those who are committed to the norm of voting may feel that there is a stigma attached to reporting nonvoting to the interviewer, whereas those who are not committed to this norm are willing to report their nonvoting openly. The effect of these commitments is not substantially altered by the presence of third persons during the interview.

---

[11] There is some variation in overreporting related to who was present during the interview among nonvoters who said they expected to vote. But caution must be exercised not to overinterpret these apparent differences, especially given that the relationship between who was present and the reporting of voting by validated nonvoters differed sharply between 1978 and 1980.

[12] There is evidence that the preelection interview itself contributes to voting (Clausen, 1968–69); Traugott and Katosh, 1979), but this does not mean that the preelection statement of intention to vote contributes to the tendency of respondents in the postelection survey to claim falsely to have voted. As the 1978 vote validation study shows, there is substantial overreporting even when there is no preelection interview. Moreover, we have indirect evidence that the preelection statement that one plans to vote does not lead to an increase in the tendency to report that one voted after the election. In 1980, as in other presidential year surveys, most respondents were questioned by the same interviewer in both the pre- and postelection interview. If respondents falsely report voting in order to appear consistent, we would expect false reporting to be more frequent among those who were interviewed by the same interviewer in the pre- and postelection surveys. We compared the extent of false reports of voting among two subsets of validated nonvoters who said that they expected to vote: those who were questioned by the same interviewer and those who were questioned by different interviewers. We found no significant difference in reported turnout between these two subsets of validated nonvoters.

These results have more general implications for the analysis of overreporting of voting. While previous research has suggested that such overreports are not linked to social background characteristics of respondents (other than race) or to general political attitudes, overreporting may be partly a result of a more narrowly focused attitude toward the act of voting itself or toward the institution of elections (cf. Dennis, 1970).[13] If so, it is premature to conclude that respondent characteristics other than race do not account for vote overreporting.

Conclusion

We have shown that the presence of third parties during interviews is far more frequent than is commonly assumed. However, the presence of third parties does not explain why respondents say that they voted when they did not vote. There may be a tendency for nonvoters to be somewhat more honest in the presence of children, especially when their spouse is also present. This does not alter our overall conclusion that the tendency of many respondents to give the socially approved answer is not strongly affected by the presence of others during the interview.

The respondent's stated intention to vote, however, is a powerful predictor of subsequent voting overreports. Respondents who are positively inclined toward voting but who for some reason do not actually vote are likely to report falsely that they voted; respondents who are not inclined toward voting are not only unlikely to vote but are also highly unlikely to report falsely that they voted.

The results suggest that additional efforts to understand the motivational basis of voting and nonvoting could help to account for the extent of vote overreporting. But they also suggest that additional efforts to avoid contamination of interviews by eliminating third parties are not likely to reduce the exaggeration of self-reported vote.

The latter result is consistent with Bradburn and Sudman's (1980) conclusion that the presence of third parties has a minimal effect on survey response patterns. But it does not close the question whether it is desirable in general to have third parties present during interviews. There is a growing body of research on this topic that has not received sufficient attention from survey research specialists. In fertility surveys, studies of sex roles, and the study of the family, whether to interview spouses (or other family members) separately or together for the collection of both objective and subjective data is an important problem that demographers have addressed (e.g., Hermalin, 1975; Thomson, 1978; Thomson and Williams, 1982; Casterline and Chidambaram, 1984; Koenig et al., 1984).

---

[13] This interpretation is not consistent with the reported lack of an empirical relationship between the sense of citizen duty and vote overreporting. The citizen duty items all focus on people's sense of the importance of elections. But the statement of an intention to vote ("Do you expect to vote . . .?") may be a more valid reflection of a person's attitude toward the election than are the more general citizen duty items.

A general conclusion from the demographic literature is that there are sometimes advantages to having third parties present during interviews. This study cannot point to any advantages to having third parties present. But it does suggest that having third parties present during the interview does not lead to exaggeration of reported voting.

References

Abramson, Paul R., John H. Aldrich, and David W. Rohde
1983 Change and Continuity in the 1980 Elections, rev. ed. Washington, D.C.: Congressional Quarterly Press.
Abramson, Paul R., and William Claggett
1984 "Race-related differences in self-reported and validated turnout." Journal of Politics 46:719–38.
Anderson, Barbara A., and Brian D. Silver
1986 "The validity of survey responses: interviews of multiple respondents in a household from a survey of Soviet emigrants." University of Michigan, Population Studies Center, Research Report No. 86-89 (January).
Bradburn, Norman M.
1983 "Response effects," pages 289–328 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.
Bradburn, Norman M., Seymour Sudman, and Associates
1980 Improving Interview Method and Questionnaire Design: Response Effects to Threatening Questions in Survey Research. San Francisco: Jossey-Bass.
Cahalan, Don
1968–69 "Correlates of respondent accuracy in the Denver validity survey." Public Opinion Quarterly 32:607–21.
Casterline, John B., and V. C. Chidambaram
1984 "The presence of others during the interview and the reporting of contraceptive knowledge and use," pages 267–98 in John A. Ross and Regina McNamara (eds.), Survey Analysis for the Guidance of Family Planning Programs. Liege, Belgium: ORDINA Editions.
Clausen, Aage R.
1968–9 "Response validity: vote report." Public Opinion Quarterly 32:588–606.
Dennis, Jack
1970 "Support for the institution of elections by the mass public." American Political Science Review 64:819–35.

Feldman, J. J., Herbert H. Hyman, and Clyde W. Hart
1951 "A field study of interviewer effects on the quality of survey data." Public Opinion Quarterly 15:734–61.

Hermalin, Albert I.
1975 "Models for the analysis of inconsistent reports of an event from two sources—with application to the accuracy of husband-wife reports." Presented to the Annual Meeting of the Population Association of America (Seattle, WA), April 17–19.

Hill, Kim Quaile, and Patricia A. Hurley
1984 "Nonvoters in voters' clothing: the impact of voting behavior misreporting on voting behavior research." Social Science Quarterly 65:199–206.

Jackman, Mary R., and Jackman, Robert W.
1983 Class Awareness in America. Berkeley and Los Angeles: University of California Press.

Katosh, John P., and Michael W. Traugott
1981 "The consequences of validated and self-reported voting measures." Public Opinion Quarterly 45:519–35.

Koenig, M. A., G. B. Simmons, and B. D. Misra
1984 "Husband-wife inconsistencies in contraceptive use responses." Population Studies 38:281–98.

Menneer, Peter
1979 "Retrospective data in survey research," pages 142–53 in Louis Moss and Harvey Goldstein (eds.), The Recall Method in Social Surveys. University of London Institute of Education, Studies in Education, new series 9.

Sigelman, Lee
1982 "The nonvoting voter in voting research." American Journal of Political Science 26:47–56.

Sudman, Seymour, and Norman M. Bradburn
1974 Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine.

Thomson, Elizabeth
1978 "The surrogate respondent: consistency of wives' and husbands' responses in a fertility survey and replication. Unpublished manuscript, Department of Sociology, University of Wisconsin.

Thomson, Elizabeth, and Richard Williams
1982 "Beyond wives' family sociology: a method for analyzing couple data." Journal of Marriage and the Family 44:999–1008.

Traugott, Michael W., and John P. Katosh
1979 "Response validity in surveys of voting behavior." Public Opinion Quarterly 43:359–77.

Verba, Sidney, and Norman H. Nie
1972 Participation in America: Political Democracy and Social Equality. New York: Harper and Row.

Weiss, Carol H.
1968–9 "Validity of welfare mothers' interview responses." Public Opinion Quarterly 32:622–33.

Opinions on Fictitious Issues: The Pressure to Answer Survey Questions

GEORGE F. BISHOP, ALFRED J. TUCHFARBER
AND ROBERT W. OLDENDICK

A VARIETY of studies since the early years of public opinion research have demonstrated that people are quite willing to express opinions on fictitious issues, objects, and events.[1] Bishop and his associates (1980,

---

[1] For an early example, see Hartley's (1946) studies of prejudice in which large majorities of college students were willing to express opinions about fictitious nationalities (e.g., "Wallonians"). See also Kolson and Green's (1970) demonstration of opinions among grade-school children about a nonexistent political figure (Thomas Walker) and Collett and O'Shea's (1976) study of people giving directions to a fictional place. References are frequently made, too, to Gill's (1947) report of large numbers of people who were willing to offer opinions on a fictional "Metallic Metals Act," but Schuman and Presser (1981) have pointed out that this report is just an anecdote with no documentation provided on the population or the sample size.

Abstract This research shows, as have previous studies, that a substantial number of people will offer opinions on fictitious topics in the context of a survey interview. The results of a split-ballot experiment (and a replication) with three fictitious public affairs issues suggest that people give opinions on such matters in large part because of the pressure to answer survey questions that is created by the way in which the questions are asked and by the manner in which "don't know" responses are handled by the interviewer. The findings also indicate that there may be little or no relationship between an individual's willingness to admit ignorance and his or her inclination to volunteer opinions on fictitious issues, as measured by at least one item from the Marlowe-Crowne Social Desirability Scale. In addition, the data demonstrate that blacks and less well-educated respondents are more likely than their counterparts to offer opinions on fictitious issues.

George F. Bishop is Associate Professor of Political Science and a Senior Research Associate in the Institute for Policy Research at the University of Cincinnati. Alfred J. Tuchfarber is Associate Professor of Political Science and Director of the Institute for Policy Research. Robert W. Oldendick is Assistant Director of the Institute for Policy Research. This research project was supported by a grant from the Procter & Gamble Company of Cincinnati, Ohio. The authors would like to thank Jack Henry, Manager of Marketing Research at P&G, for funding the project. The authors would also like to thank the anonymous referees for their helpful comments on an earlier version of the manuscript.

Public Opinion Quarterly Vol. 50:240–250 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-240/$2.50

1983), for example, have found that a sizable percentage of respondents (roughly 30 percent) will give an opinion on a fictitious "1975 Public Affairs Act" if they are not offered an explicit response alternative which allows them to admit that they "don't know" anything about it. Similarly, Schuman and Presser (1981) have discovered that about 30 percent of the respondents in their national surveys will, if not given an explicit chance to say thay have "no opinion" on the subject, answer questions about real, but highly obscure, pieces of legislation such as The Agricultural Trade Act of 1978 and The Monetary Control Bill of 1979. Such are the pressures in a typical interview to answer survey questions.

From such studies, we have also learned that some people are more likely than others to offer opinions on obscure or fictitious issues: blacks more likely than whites, and the less educated more likely than the better educated (Bishop et al., 1980; Schuman and Presser, 1981; Ch. 5). The latter finding is somewhat puzzling because, as Schuman and Presser have noted, less educated respondents are generally *less likely,* not *more likely,* to volunteer opinions on public affairs issues. What explains the difference? Schuman and Presser have suggested that better educated people may be more willing than those who are less well-educated to admit their ignorance on topics which they do not recognize. Bishop and his colleagues (1980) have proposed a similar explanation, hypothesizing that blacks and less educated respondents, of whatever race, may offer opinions on the Public Affairs Act to "save face" with "our presumptively white middle-class interviewing staff . . . calling from the University. . . ."[2] Common to both these explanations is the idea that people have a disposition or personality characteristic which makes them either more or less willing to admit their ignorance on public affairs.

Neither of these dispositional hypotheses, however, has been directly tested or pitted against an equally plausible explanation, also suggested by Schuman and Presser: "that low-educated respondents are especially susceptible to pressures in the interview to give an opinion even where they have none (p.152)." Furthermore, Schuman and Presser have noticed that, in some instances, interviewers increase the pressure on respondents to give an answer by (inappropriately) repeating the question to them even though they have already indicated that they "don't know" anything about the subject. Respondents thus get the message that the interviewer *expects* them to answer the question. The majority of respondents do not conform to this implicit demand, but a sizable number do, indicating that the pressure to give an opinion in this situation can be quite compelling. That, in fact, is what the experiment described here was designed to demonstrate: that people give opinions on obscure or fictitious issues largely because of the "demand characteristics" of the survey interview (cf. Orne, 1962). The experiment was also designed to test the alternative dispositional hypothesis: that people give opinions on such matters because they are unwilling to admit that they are ignorant. Included in the design, as well, was an analysis of the response effects of the two major demographic variables in previous research: race and education.

---

[2] Bishop et al. (1980, fn. 10) report an almost perfect correlation (.99) between the race of the respondent perceived by the interviewer in the first wave of a telephone panel study and the respondent's self-reported race in the second wave, indicating that the race of the respondent (and by implication, the interviewer's) can be accurately inferred over the phone.

Research Design

In November and early December of 1983 we conducted the first of two split-ballot experiments on people's responses to questions about three fictitious issues: The 1975 Public Affairs Act, The Agricultural Trade Act of 1983, and The Monetary Control Bill of 1983. The questions were all asked, together, near the middle of the interview, as part of a multipurpose, random-digit-dialed telephone survey,[3] and either immediately before or after questions about three public affairs issues which were not fictitious (see Appendix). The latter items were included to provide a cover for asking respondents about the fictitious issues as well as a test and control for the influence of question context.[4]

The question about the Public Affairs Act was drawn from the experiments by Bishop et al. (1980). The questions about the Agricultural Trade Act and the Monetary Control Bill were adapted from the items developed originally by Schuman and Presser (1981: Ch. 5): The Agricultural Trade Act of *1978* and The Monetary Control Bill of *1979* (emphais added). We changed the year of these legislative acts—to

---

[3] The response rate (fully completed interviews) for this Greater Cincinnati Survey was 71.5 percent. The refusal rate was 15.8 percent. The remainder consisted of partially completed interviews (3.2 percent) and potential interviews that were not completed because of a language barrier, a hearing problem, illiteracy, senility, or physical illness (5.6 percent) or because the selected respondent was away on vacation, a business trip, or avoided an appointment for some other reason (3.9 percent).

[4] The question about whether the federal government is doing too many things that should be handled by the private sector was adapted from an item developed originally by the National Opinion Research Center (see NORC Study 4179, December 1973; 14, University of Chicago). The questions on trade with the Soviet Union and government spending were both adapted from items designed by the Center for Political Studies at the University of Michigan (see the 1982 National Election Study Cross-Section Interview: 52 and 55).

1983—to make them appear up-to-date and thus more compelling as issues on which the respondent should probably have an opinion.[5] To our knowledge, there were no such legislative matters pending in Congress at the time of survey. Nor is it likely that any respondent knew anything about the real, but obscure, legislation from which these items were initially derived.

QUESTION FORMS

The questions were asked in one of three forms, to which respondents were randomly assigned:

1. *Form A*. In this form the questions were asked with an explicit response alternative (a filter) which allowed respondents to indicate that they hadn't thought much about the issue (see Appendix).[6]

2. *Form B*. Here the questions were asked without an explicit filter. But if the respondents volunteered that they didn't know anything about the topic or gave a qualified answer to the question other than the two response alternatives that were read to them (e.g., agree/disagree), the interviewers were instructed *not to probe* them in any way, but rather to record, verbatim, whatever the respondent said.

3. *Form C*. The questions in this form were also asked without a filter. However, if the respondents volunteered a "don't know" response or offered a qualified answer to the question, the interviewers were told to pressure them into selecting one of the two forced-choices (e.g., favor/oppose) by repeating the question, by rereading the choices, or both.

Forms B and C were varied to test Schuman and Presser's proposition about the pressure interviewers put on respondents by probing, or what we call the "demand characteristics" hypothesis.

QUESTION ORDER AND CONTEXT

In addition, we varied the order in which the groups of questions were asked, randomly assigning respondents to one of two conditions: in one condition, respondents were asked the group of questions about the fictitious issues immediately *before* the group of questions about the nonfictitious issues; in the other condition, immediately *after* them. Within each of these groups, however, the three questions were asked in the same sequence in both conditions (see Appendix). This manipulation allowed us to test whether respondents would be more susceptible to being pressured into answering the questions about the fictitious issues when they were asked in one context rather than another.

---

[5] The wording of the filter used with our adaptation of these questions—i.e., "or haven't you thought much about this issue?"—differs from the one used by Schuman and Presser—i.e., "or do you not have an opinion on that issue?" This variation may have had some effect on the proportion of DK responses observed in our experiment, but it is probably a minor one (see Bishop et al., 1983 on the effects of differently worded filter questions).

[6] Respondents receiving this form were *not* probed if they volunteered a "don't know" response or gave a qualified answer to the item. This, we reasoned, was consistent with the intent of the filtered form, which was to make it acceptable not to have an opinion on an issue. It also seemed consistent not to pressure respondents who received this form into selecting one of the forced-choices, but rather to let them "freely" qualify their answers if they wanted to.

MEASUREMENT OF WILLINGNESS TO ADMIT IGNORANCE

To tap the disposition to admit ignorance, we adapted an item from the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1964) and asked it along with four other items from the same scale which allowed us to construct a general index of the tendency of people to present themselves to others in a favorable light. The five items, the fourth of which measured a person's willingness to admit ignorance, were worded as follows.[7]

> Now I'd like to read you some statements about personality characteristics. As I read each one, please tell me whether the statement is true or false as it applies to you personally.
>
> 1. I have never deliberately said something that hurt someone's feelings—is that true or false?
> 2. I'm always willing to admit it when I make a mistake—is that true or false?
> 3. No matter who I am talking to, I'm always a good listener—is that true or false?
> 4. When I don't know something I don't at all mind admitting it—is that true or false?
> 5. I'm always courteous, even to people who are disagreeable—is that true or false?

Because responses to the question about admitting ignorance might influence responses to questions about the fictitious issues, or vice versa, we decided to randomize the location of the Marlowe-Crowne items: they were asked shortly before the fictitious and nonfictitious issues in one condition; shortly after them, in the other. To hide the connection between the question about admitting ignorance and the fictitious issues, we not only embedded it among the other Marlowe-Crowne items but also created a brief "buffer" between these items and the issues with two questions about the local major league baseball team. The latter were asked immediately *after* the Marlowe-Crowne items, and just *before* the issue questions, in one condition; immediately *before*, and just *after*, in the other.

---

[7] If respondents tried to qualify their answers to these questions by saying such things as, "Well that's true sometimes, but not always," interviewers were instructed to probe them by asking: "Is it *mostly true* or *mostly false*?" In addition, they were instructed to probe all "don't know" responses by repeating the question. This was done to minimize missing data.

REPLICATION

In May and early June of 1984 we replicated the entire experiment, exactly, again as part of a multipurpose, random-digit-dialed telephone survey.[8]

Findings

The data in Table 1 demonstrate that the form in which the questions are asked, and the way in which the "don't know" (DK) responses are dealt with, makes a substantial difference in the results.[9] People were much more likely to offer an opinion on a fictitious issue when the question was asked without an explicit filter (Forms B and C) than when it was asked with one (Form A). And they were even more likely to offer an opinion on the unfiltered form when "don't know" responses were probed (Form C) than when they were not (Form B). This pattern occurred on all three of the issues, and in both surveys.[10]

Notice, however, that respondents were generally more likely to offer opinions on both the Agricultural Trade Act and the Monetary Control Bill than they were on the Public Affairs Act, probably because they were more likely to confuse the former topics with subjects with which they were familiar, such as the questions we asked about trade with the Soviet Union and government spending (see Appendix). People can thus be more easily manipulated into giving opinions on subjects which seem at least faintly familiar than those which do not.

EFFECTS OF WILLINGNESS TO ADMIT IGNORANCE

We found little or no relationship between an individual's willingness to admit ignorance and his or her tendency to offer opinions on

---

[8] The response rate for this Greater Cincinnati Survey was 71.3 percent. The refusal rate was 14.8 percent. The rest consisted of partially completed interviews (2.8 percent) and potential interviews that were not completed because of a language barrier, a hearing problem, illiteracy, senility, or physical illness (5.6 percent) or because the selected respondent was away on vacation, a business trip, or avoided an appointment for some other reason (5.5 percent).

[9] Because there were no significant differences in responses to the fictitious issues, in either of the two surveys, when the questions were asked before or after the nonfictitious issues (data not shown here), we have combined the data for these two context conditions in constructing Table 1.

[10] The differences between Forms B and C tend to be somewhat greater in the May/June 1984 survey, especially on the Public Affairs Act, indicating that the interviewers may have pressured respondents more heavily and consistently, perhaps because of greater emphasis in training for this survey on how to probe respondents who give "don't know" responses.

Table 1. Percent with Opinion on Public Affairs Act, Agricultural Trade Act, and Monetary Control Bill by Question Form and Survey

	Filtered	*Unfiltered*	
	Form A (DK's Not Probed)	*Form B (DK's Not Probed)*	*Form C (DK's Probed)*
Public Affairs Act	3.7	24.3	31.5
(Nov./Dec., 1983)[a]	(408)	(395)	(399)
Public Affairs Act	3.6	22.2	42.3
(May/June, 1984)[b]	(397)	(411)	(402)
Agricultural Trade Act	12.1	38.9	49.0
(Nov./Dec., 1983)[c]	(405)	(399)	(398)
Agricultural Trade Act	10.5	35.5	55.9
(May/June, 1984)[d]	(398)	(411)	(403)
Monetary Control Bill	13.2	33.7	46.3
(Nov./Dec., 1983)[e]	(408)	(397)	(397)
Monetary Control Bill	14.3	38.4	54.7
(May/June, 1984)[f]	(398)	(411)	(404)

[a] Form A vs. B vs. C: $\chi^2=106.05$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=4.86$, df=1, $p<.03$).

[b] Form A vs. B vs. C: $\chi^2=171.01$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=36.93$, df=1, $p<.0001$).

[c] Form A vs. B vs. C: $\chi^2=130.21$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=7.91$, df=1, $p<.01$).

[d] Form A vs. B vs. C: $\chi^2=183.33$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=33.02$, df=1, $p<.0001$).

[e] Form A vs. B vs. C: $\chi^2=104.05$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=12.57$, df=1, $p<.001$).

[f] Form A vs. B vs. C: $\chi^2=143.68$, df=2, $p<.0001$ (Form B vs. C: $\chi^2=21.00$, df=1, $p<.0001$).

fictitious issues, regardless of the form in which the questions were asked (data not shown here).[11] The failure to find such a relationship does not, however, mean that the dispositional hypothesis can be discarded because the measure we used was admittedly a rather weak one, based on a single item, the responses to which were highly skewed: over 90 percent of the respondents in the first survey, and nearly 90 in the second, said they don't at all mind admitting it when they "don't know" something. So there was little or no variance in the item, and therefore little opportunity for it to differentiate between respondents who give opinions on fictitious issues and those who do not. Perhaps, then, a better measure, or set of measures, of a person's willingness to admit ignorance would result in a different conclusion.

---

[11] Nor did we find any significant relationship, in either study, between a respondent's score on an overall index of social desirability—formed by summing responses to all five of the Marlowe-Crowne items—and his or her response to any of the fictitious issues (data not shown).

Table 2. Percent with Opinion on Public Affairs Act, Agricultural Trade Act, and the Monetary Control Bill by Race, by Education, by Question Form (Fall 1983 and Spring 1984 Surveys Combined)

	Filtered Form (A)				*Unfiltered Forms (B & C Combined)*			
	Better Educated		*Less Well Educated*		*Better Educated*		*Less Well Educated*	
	Blacks	*Whites*	*Blacks*	*Whites*	*Blacks*	*Whites*	*Blacks*	*Whites*
Public Affairs Act[a]	4.0	5.0	4.2	2.2	41.5	22.1	52.2	30.9
	(50)	(299)	(71)	(368)	(82)	(633)	(159)	(706)
Agricultural Trade Act[b]	16.0	11.4	4.3	12.2	51.8	38.4	59.0	46.5
	(50)	(297)	(70)	(369)	(83)	(635)	(156)	(710)
Monetary Control Bill[c]	16.0	18.5	16.9	9.5	56.6	40.8	59.5	40.5
	(50)	(298)	(71)	(370)	(83)	(633)	(158)	(709)

NOTE: Better educated = some college or more; less well-educated = high school education or less.

[a] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 13.72$, df = 1, $p < .001$); less well-educated blacks vs. whites ($\chi^2 = 25.07$, df = 1, $p < .0001$).

[b] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 4.93$, df = 1, $p < .03$); less well-educated blacks vs. whites ($\chi^2 = 7.50$, df = 1, $p < .01$).

[c] Differences on filtered form *not* statistically significant. Differences on unfiltered forms: better educated blacks vs. whites ($\chi^2 = 6.92$, df = 1, $p < .01$); less well-educated blacks vs. whites ($\chi^2 = 18.20$, df = 1, $p < .0001$).

EFFECTS OF RACE AND EDUCATION

As expected from previous research, both race and education had a significant influence on responses to the fictitious issues, depending upon the form in which the questions were asked (see Table 2).[12] On the filtered form (A), the differences between blacks and whites, and between the better educated and the less well educated, were not statistically significant on any of the three times. Most respondents, in fact, regardless of their race or education, said they hadn't thought much about any of the fictitious issues when asked about them. On the unfiltered form, however, the differences between groups were sizable as well as statistically significant. Blacks offered opinions on all three issues significantly more often than whites, whether they were better educated or not. Notice, too, that education had an independent effect, particularly on responses to the Public Affairs Act: the less well educated offered opinions on the issues more often than the better educated, whether they were black or white. But the effects of education were not as strong as those of race; in fact, on the Monetary Control Bill, there was no significant difference in response by education when race was held constant. Race is thus the more important of the two.[13]

---

[12] Because of small cell sizes in the black subsample, the data for the two surveys have been combined and education has been collapsed into two categories, an admittedly crude control, but inescapable under the circumstances.

Conclusion

The findings suggest that people give opinions on fictitious issues, in large part, because of the *pressure to answer* which is created by the form in which the questions are asked and the manner in which the "don't know" responses are handled by the interviewer. The data also indicate that people can be more easily pressured into giving an opinion on a fictitious issue when the topic seems familiar to them (e.g., agricultural trade), probably because they confuse it with a related subject on which they do have an opinion. The more a person knows about a subject, however (e.g., politics, sports), the less likely he or she is to make such a mistake; in fact, the more they know about the subject, the more easily they can recognize what is familiar and what is not (e.g., a fictional politician or baseball team). Blacks and less well-educated respondents may thus be more likely to offer opinions on fictitious issues because they are generally less well informed about politics and therefore less able to tell whether the topic is something that "belongs."[14] As a general principle, then, we would hypothesize that the less knowledgeable a person is about a given subject, the more easily he or she can be confused and pressured to give an opinion about it. Such are the "demand characteristics" of the survey interview.

What are the implications of these findings for the practice of probing? Surely we should not pressure people to give opinions when they plead ignorance. Most of the problem can, in fact, be avoided by giving respondents an explicit opportunity to say they have "no opinion" on the subject—for example, by using a filter question. But on many public affairs issues this will mean losing 20 to 25 percent of the cases for the analysis (see Bishop et al., 1983; Schuman and Presser, 1981:

---

[13] We also discovered that, with the exception of the Public Affairs Act, men were somewhat more likely than women to offer opinions on the fictitious issues on both the filtered and unfiltered forms, but the differences were relatively small—less than 10 percent on the average (data not shown). Younger people were significantly more likely than older people to give opinions on the Monetary Control Bill when the question was asked in unfiltered form, but other than that there were no significant differences among age groups.

[14] We have found in a previous survey, for example, that blacks are significantly less likely than whites to know the names of their United States senators, the length of a senator's term of office, and which political party has the most members in the U.S. House of Representatives (data not presented here).

Ch. 4). That, for many pollsters, however, may be too big a price to pay in the loss of representativeness. Nor does it make good copy to say that many people have no opinion on the issues of the day. The alternative is to create a spurious form of representativeness by minimizing the "don't knows." Perhaps we should stop probing altogether, except to ask respondents what they mean when they say "I don't know." For probing presumes that the respondent has an opinion on the issue that merely needs to be drawn out. Yet, how can we ever know that in advance? Is it not time, then, that we reexamined the practice of probing in public opinion surveys? And with it, the meaning of what we so routinely record as "DK."

Appendix

WORDING OF THE QUESTIONS ON FICTITIOUS AND NONFICTITIOUS ISSUES

Form A	*Forms B and C*
Some people feel that The 1975 Public Affairs Act should be repealed—do you agree or disagree with this idea, or haven't you thought much about this issue?	Some people feel that *The 1975 Public Affairs Act* should be repealed—do you agree or disagree with this idea?
Congress has been considering *The Agricultural Trade Act of 1984*—do you favor or oppose the passage of this act, or haven't you thought much about this issue?	Congress has been considering *The Agricultural Trade Act of 1984*—do you favor or oppose the passage of this act?
Congress has also been considering *The Monetary Control Bill of 1984*—do you favor or oppose the passage of this bill, or haven't you thought much about this issue?	Congress has also been considering *The Monetary Control Bill of 1984*—do you favor or oppose the passage of this bill?
. . . Do you think the government is trying to do too many things that should be left to individuals and private businesses, do you think the government should do even more to solve our country's problems, or haven't you thought much about this issue?	. . . Do you think the government is trying to do too many things that should be left to individuals and private businesses, or do you think the government should do even more to solve our country's problems?
. . . Do you think the United States should do more trading and business with the Soviet Union, do you feel that the United States should cut back on its economic ties with the Soviet Union, or haven't you thought much about this issue?	. . . Do you think the United States should do more trading and business with the Soviet Union, or do you feel that the United States should cut back on its economic ties with the Soviet Union?

Form A (cont.)	*Forms B and C* (cont.)
. . . Do you think the federal government should provide fewer services, even in areas such as health and education, in order to reduce spending, do you feel it is important for the government to provide many more services even if it means an increase in spending, or haven't you thought much about this issue?	. . . Do you think the federal government should provide fewer services, even in areas such as health and education, in order to reduce spending, or do you feel it is important for the government to provide many more services even if it means an increase in spending?

References

Bishop, George F., Robert W. Oldendick, and Alfred J. Tuchfarber
1983 "Effects of filter questions in public opinion surveys." Public Opinion Quarterly 47:528–46.

Bishop, George F., Robert W. Oldendick, Alfred J. Tuchfarber, and Stephen E. Bennett
1980 "Pseudo-opinions on public affairs." Public Opinion Quarterly 44:198–209.

Collett, Peter, and Gregory O'Shea
1976 "Pointing the way to a fictional place: a study of direction giving in Iran and England. European Journal of Social Psychology 6:447–58.

Crowne, Douglas P., and David Marlowe
1964 The Approval Motive. New York: Wiley.

Gill, Sam
1947 "How do you stand on sin?" Tide magazine, March 14, pp.72 ff.

Hartley, Eugene L.
1946 Problems in Prejudice. New York: Octagon Press.

Kolson, Kenneth L., and Justin J. Green
1970 "Response set bias and political socialization research." Social Science Quarterly 51:527–38.

Orne, Martin J.
1962 "On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications." American Psychologist 17:776–83.

Schuman, Howard, and Stanley Presser
1981 Questions and Answers in Attitude Surveys. New York: Academic Press.

Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys

ROBERT M. GROVES AND LOU J. MAGILAVY

INTERVIEWER variance or interviewer effects reflect the tendency for answers provided by the respondent and recorded in a questionnaire to vary depending on which interviewer is assigned to the respondent. In order to obtain estimates of these effects, "interpenetrated" designs are typically introduced, which randomly assign sample cases to interviewers. Although several researchers have touted the possibility of cheap interpenetration in centralized telephone surveys (see Groves and Kahn, 1979; Tucker, 1983), routine measurement has been found to be a greater challenge than previously assumed. This article sum-

---

Abstract Estimates of interviewer effects on survey statistics are examined from nine surveys conducted over a six-year period at the Survey Research Center. Estimates of intraclass correlations associated with interviewers are found to be unstable, given the number of interviewers (30–40) used on most surveys. This finding calls into question inference from earlier studies of interviewer effects. To obtain more reliable information about magnitudes of interviewer effects, generalized effects are constructed by cumulating estimates over statistics and surveys. These generalized correlations are found to be somewhat smaller than those reported in the past literature. Few differences in generalized interviewer effect measures are found between open and closed questions or between factual and attitudinal questions. Small reductions in effects *were* obtained when a Computer Assisted Telephone Interviewing (CATI) system was used; there was some evidence of elderly respondents being more susceptible to interviewer effects; the number and type of second responses to open questions were affected by interviewer behavior; and changes in interviewing techniques reduced interviewer effects.

Robert M. Groves is an Associate Research Scientist in the Survey Research Center and an Associate Professor in the Department of Sociology at The University of Michigan. Lou J. Magilavy is a Ph.D. candidate in the Department of Biostatistics at The University of Michigan. Data in this article were collected in projects supported by the National Science Foundation and the National Center for Health Statistics. The views in this article do not necessarily represent those of either organization. An earlier version of this article was presented at the meetings of the International Statistical Institute, 1983. The authors appreciate the comments of Leslie Kish, Stanley Presser, and Howard Schuman on earlier versions, and extend their gratitude to Richard Curtin for access to the data from the Surveys of Consumer Attitudes.

Public Opinion Quarterly Vol. 50:251–266 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-251/$2.50

marizes six years of research in the measurement of interviewer effects in centralized telephone surveys, on topics ranging from health experiences to political attitudes. We attempt to address two weaknesses in past research—problems of estimation (obtaining precise and accurate estimates of interviewer effects) and problems of explanation (learning what design features affect the magnitude of interviewer effects).

Study Designs

Eight of the nine surveys examined used a two-stage stratified national sample of randomly generated telephone numbers, following the design of Waksberg (1978). The ninth survey, the 1980 Post Election Study, was a telephone reinterview with respondents chosen earlier in an area probability sample. In all surveys one adult was randomly selected in each sample household. As shown in Table 1, response rates ranged from 59 percent to 87 percent, with unanswered numbers included in the denominators of these rates.[1] Row 4 lists the number of interviews obtained from the random assignment of telephone numbers to interviewers. Interviews resulting from nonrandom assignments (e.g., refusal conversions, select appointments) were excluded from both those numbers and the following analyses. Also presented is the number of interviewers employed for each study and the average workload (i.e., number of interviews) per interviewer. Not only does average workload vary across studies, from 12 for the November 1981 Survey of Consumer Attitudes to 58 for the Health in America Survey, but workloads within studies exhibit a similar degree of variation. All the studies collected standard demographic data in addition to data on the specific topics listed in Table 1. The 30 variables analyzed for the Survey of Consumer Attitudes are identical in each of the five months. The number of statistics analyzed for each study ranges from 25 to 55.

Estimation of Interviewer Effects

THE INTERVIEWER EFFECTS ESTIMATOR

To compare interviewer effects across statistics and to maximize comparability to past interviewer variance research we utilize an intraclass correlation, based on a one-way analysis of variance, using interviewers as the factor, $\rho^{*}_{int} = S_a^2/(S_a^2 + S_b^2)$, where S_a^2 is the between interviewer variance and S_b^2 is the within interviewer variance, with

---

[1] The lowest response rate, 59 percent, reported for the Study of Telephone Methodology reflects an exceptionally large proportion of unanswered numbers. In later studies, improved methods of identifying nonworking numbers have caused a decrease in this proportion and consequently a more accurate response rate.

Table 1. Study Descriptions and Summary of Overall Findings

Study Description and Summary Statistics	*Study of Telephone Methodology*	*Health and Television Viewing*	*Health in America*	*1980 Post Election Study*	*Monthly Survey of Consumer Attitudes*				
					Nov '81	*Dec '81*	*Jan '82*	*Feb '82*	*Mar '82*
Field period	Apr–May 1976	Mar–Apr 1979	Oct–Dec 1979	Nov–Dec 1980	Nov '81	Dec '81	Jan '82	Feb '82	Mar '82
Response rate	59%	67%	80%	87.3%	73.5%	72.8%	73.3%	76.9%	73.8%
No. interviewers	37	30	33	22	31	28	21	25	26
No. interviews	1529	954	1918	697	370	350	386	379	366
Average workload	41.3	31.8	58.1	31.7	11.9	12.5	18.4	15.2	14.1
Major topics	Political economic & social issues	Health and TV viewing	Health	Political issues	Economic issues	Economic issues	Economic issues	Economic issues	Economic issues
No. variables analyzed	25	55	25	42	30	30	30	30	30
Range of ρ^*_{int}	$-.0080$ $.0560$	$-.0150$ $.1650$	$-.0070$ $.0097$	$-.0154$ $.1710$	$-.0217$ $.0895$	$-.0373$ $.0546$	$-.0221$ $.0916$	$-.0419$ $.0657$	$-.0356$ $.0729$
Mean values									
ρ^*_{int}	.0089	.0074	.0018	.0086	.0184	.0057	.0163	.0090	.0067
$deff_{int}$	1.36	1.23	1.10	1.26	1.20	1.07	1.28	1.13	1.09

large positive values of ρ^*_{int} reflecting large interviewer effects. This is the measure described more fully by Kish (1962) and used frequently in this literature. His studies of factory workers' job attitudes produced ρ^*_{int}'s of $-.03$ to .09 (mean = .02). O'Muircheartaigh and Marckwardt (1980) presented fertility statistics most susceptible to interviewers, with ρ^*_{int}'s between .00 and .19 (mean = .07). Feather (1973) in a health survey found a range of ρ^*_{int}'s of $-.01$ to .03 (mean = .006). Freeman and Butler (1976), in a study of attitudes toward mental retardation, displayed unusually high values of ρ^*_{int} (mean = .04), that might be attributable to the use of nonprofessional interviewers. The overall average for all these statistics from personal interview surveys is .018. Tucker (1983) presented mean ρ^*_{int}'s from 11 telephone surveys ranging from $-.003$ to .008 (overall mean = .004) for political and social attitudes.[2]

For this research we estimated ρ^*_{int} after interpenetrating the sample within shifts of a centralized telephone interviewing facility. This randomization of assignment within shift is an added constraint on the model, the effect of which was judged to be negligible in these studies.[3]

It is important to realize that the effect of the interviewer on the overall variance of survey statistics is a function of ρ^*_{int} *and* the size of the interviewer workload, m. The variance of a survey statistic is inflated by the factor $deff_{int} = [1 + \rho^*_{int}(m - 1)]$. Thus, if interviewers complete large numbers of cases, say 50, even ρ^*_{int}'s as small as .01 can inflate the variance by a factor of 1.5 over that expected without interviewer effects.[4]

---

[2] In citing previously published values of ρ^*_{int}, we note that most of these accounts provide poor descriptions of the survey design features that might affect interviewer errors—the question format and wording, interviewer training procedures, nonresponse and recording error differences across interviewers, and any other effects of the survey organization. Often only a small number of survey statistics are presented, sometimes only to illustrate the *largest* interviewer effects observed.

[3] A separate appendix describing interpenetration problems in a centralized facility is available from the authors.

Perhaps the most important weakness of the past research is the lack of attention to the stability of the ρ^*_{int} estimates. None of the studies above give guidance about whether similar effects might be expected in replications of the surveys. Given the model of interviewer effects, the precision of the ρ^*_{int} is a function of the number of interviewers used and the number of cases completed by each. Many of these studies used only a few interviewers, but the reader is not cautioned about the likelihood of different levels of interviewer effects being observed in replications of the survey.[5]

MAGNITUDES AND PRECISION OF ESTIMATED INTERVIEWER EFFECTS

In the space allowed it is impossible to describe the results of each study in detail, but Table 1 presents the range of ρ^*_{int} estimates obtained in each study, along with their mean values.[6] The average values of the ρ^*_{int}'s seem to lie below .01 over the nine surveys and to exhibit large variability over the different statistics examined. The highest average ρ^*_{int} (.0184) is found for one of the Consumer Attitude Surveys, but comparison over the other months for the same statistics shows means as low as .0057. The lowest average ρ^*_{int} (.0018) is found for a mainly factual survey on health conditions, Health in America. This survey also had the largest interviewer workload.[7]

Table 2 illustrates the importance of past studies' failure to measure the stability of ρ^*_{int} by presenting standard errors for 30 ρ^*_{int} statistics, 10 from each of three studies.[8] These standard errors are estimates of the dispersion of the distribution of ρ^*_{int}'s over replications of the surveys, with variation due to the selection of different interviewers from the population of all interviewers and from sampling variability of the respondents. This table illustrates well the range and magnitude of ρ^*_{int} values estimated throughout the nine surveys. *The standard errors for the measures of interviewer effects are often larger in size than the estimates themselves,* leading one to conclude that the values of ρ^*_{int} presented here could have been much larger *or much smaller* had different sets of interviewers been chosen to work on the surveys.

---

[4] This also implies that the inflation of variance is smaller for subclass statistics in the same survey, where m is smaller.

[5] Some authors base their conclusion that serious interviewer effects are present on the results of a standard F-test. This test is subject to several assumptions which may be violated. For example, the interviewer workloads often exhibit a great deal of variation and the variances within interviewer assignments may not be constant across interviewers.

[6] Some of the statistics are means computed on continuous variables. For discrete variables, ρ^*_{int} for the proportion choosing the modal response category is presented.

[7] This fact, together with other observations, has made us begin to consider enhancing the response error model underlying ρ^*_{int} to reflect larger interviewer variability during the initial cases completed by interviewers.

[8] In order to estimate the sampling variance of the ρ^*_{int}'s a jackknife estimator was used by dividing the sample into subsamples associated with each interviewer. Although the properties of the jackknife estimator are known in the case of a linear statistic, it has also performed well empirically for a variety of nonlinear estimators including regression and correlation coefficients (Kish and Frankel, 1972). These empirical results suggested its use in this case.

In order to relate our results to past research, these standard error estimates can be used to illustrate the largest ρ^*_{int}'s likely to be encountered with these designs. If we assume that the ρ^*_{int}'s are distributed roughly according to a t-distribution, we can estimate the value beyond which only 10 percent of the ρ^*_{int} values lie. These are called the "maximum" values of ρ^*_{int} in Table 2. The values of ρ^*_{int} presented in Table 1 are lower than the averages from past personal interview surveys; the *maximum values* of ρ^*_{int}'s, however, are in the general range of those estimates.

It has been assumed for some time that the measurement of interviewer effects would be enhanced by centralized telephone interviewing facilities. The ability to incorporate an interpenetrated design at a relatively small cost makes this an attractive setting for the evaluation of interviewers. However, it remains difficult to draw conclusions from the results presented here because of the imprecision of these estimates. The instability of the estimates is mainly due to the small number of interviewers employed for telephone surveys. Since the past literature in the area has generally failed to address the problem of imprecision, it is possible that many of the effects judged to be important in those studies would not be found in replications.

Remedies for Instability of Interviewer Effect Estimates

The researcher faced with unstable estimates of interviewer effects has the alternatives of ceasing their measurement, altering the survey design to improve their precision (chiefly by increasing the number of interviewers used), or increasing their precision through replication of surveys. Ceasing attempts to measure this source of survey error seems premature; the other two options deserve thorough comment.

Increasing the number of interviewers is not recommended in this case, because the hiring and training of larger numbers of interviewers may force a change in essential survey conditions that help to determine the magnitudes of interviewer effects in the survey data. If, for example, the number of interviewers were increased greatly, the quality of training or amount of supervisory attention given to each inter-

Table 2. Illustrative Values and Jackknife Estimates of Standard Errors of ρ^*_{int}

Estimated Proportions or Means	ρ^*_{int}	*S.E.*	*"Maximum"* ρ^*_{int}
HEALTH AND TELEVISION VIEWING			
Having had arthritis or rheumatism	.0227	.0132	.0400
Having had diabetes	−.0059	.0053	.0010
Not drinking beer in month	.0040	.0091	.0159
Not drinking hard liquor in month	.0152	.0117	.0305
Not visited doctor in 6 months	.0266	.0144	.0455
Reporting week of last doctor's visit	.0310	.0152	.0509
Number of ameliorative health behaviors reported	.0285	.0141	.0470
Number of ways mentioned that TV is good or bad for children	.0157	.0129	.0326
Number of "bad reactions" to medicine	.0076	.0133	.0250
Number of health symptoms reported	−.0032	.0062	.0049
HEALTH IN AMERICA			
Number of doctor visits past 2 weeks	.0092	.0076	.0192
Number of hospitalizations during the past year	.0040	.0041	.0094
Reporting excellent health status	.0085	.0077	.0186
Reporting 2–4 doctor visits past year	.0002	.0035	.0048
Mean number of chronic conditions	.0096	.0063	.0179
No bed days during the past year	−.0070	.0029	−.0032
No operations past year	−.0031	.0030	.0008
No bed days past 2 weeks	−.0057	.0025	−.0024
2 wks–6 months since doctor visit	.0003	.0042	.0058
2 wks–6 months since dental visit	.0017	.0048	.0080
SURVEY OF CONSUMER ATTITUDES, DEC '82			
Better off financially than last year	−.0232	.0171	−.0008
Saw increase in prices during the last 12 months	.0101	.0181	.0338
Percent change in prices during the last 12 months	−.0325	.0178	−.0092
Feel now a bad time to buy house	.0112	.0239	.0425
Feel now a good time to buy major household items	.0242	.0258	.0580
Expect decrease in interest rates next 12 months	.0089	.0271	.0444

Table 2. *(Continued)*

Estimated Proportions or Means	ρ^*_{int}	*S.E.*	*"Maximum"* ρ^*_{int}
Expect increase in unemployment next 12 months	−.0210	.0151	−.0012
Expect no change in their financial situation next 12 months	.0308	.0269	.0660
Expect family income to increase 15% or less next 12 months	.0546	.0385	.1050
Believe government's economic policy fair	.0508	.0361	.0981

viewer might be altered (ignoring the higher costs incurred with such a design). With departures from training guidelines interviewer variance might increase. Thus, we face the Heisenberg-like result that to measure a particular survey error well *in a single survey,* we may change its value.

Another way to obtain useful knowledge about the average levels of interviewer variability is to repeat measurement of interviewer effects over surveys. By cumulating results over surveys we hope to take advantage of the fact that averages of the same survey statistic over replications have greater stability than a single measure of it. Ideally, averages would be constructed from identical measures of the same population, using the same interviewer selection and staffing rules, and the same statistics. For example, four replications of a survey might reduce the standard errors of average ρ^*_{int}'s by half, relative to those from a single survey.[9] Rarely, however, are exactly equivalent measures obtained in several repeated studies. Lacking that, cumulation of ρ^*_{int} estimates over statistics and across surveys might be attempted within the same survey organization. This is similar to the construction of generalized sampling variance estimates in complex samples. Given the weaknesses of the other approaches, generalized interviewer effect estimates obtained by averaging over surveys are attractive.

In Figure 1 the cumulative distribution of ρ^*_{int} values is plotted for each study over all statistics for which ρ^*_{int} was calculated. It may be observed that in general the *shapes* of the distributions are related to sample size and numbers of interviewers, with the range of values of ρ^*_{int} decreasing as these increase in number. This mirrors the estimates of standard errors of ρ^*_{int}. We see that while each study yielded some

---

[9] This assumes independence in the selection of interviewers across the replications and no change in true values of interviewer effects over replications.

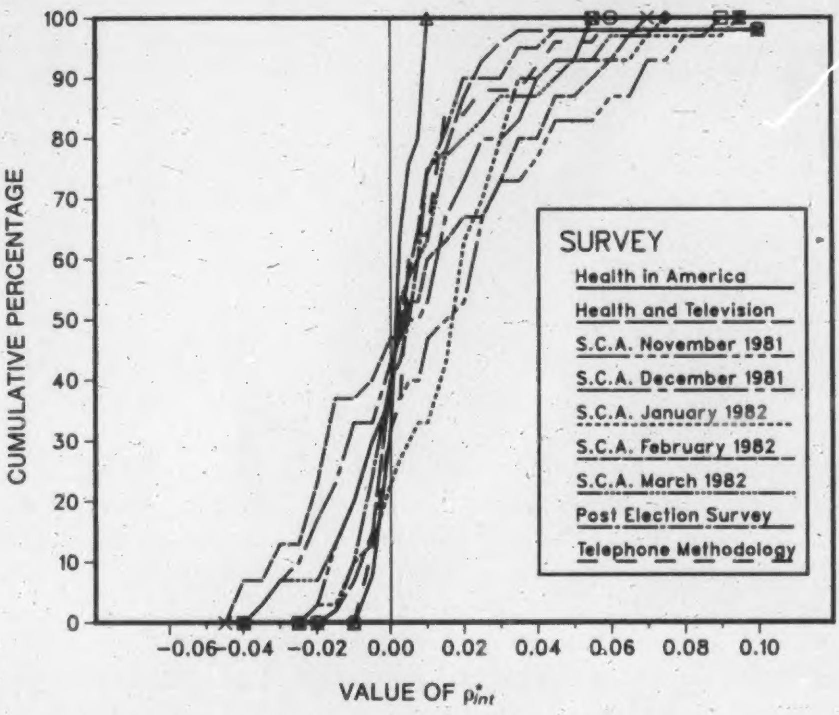


Figure 1. Cumulative Distributions of Interviewer Effect Measures (ρ^*_{int}) for Nine SRC Surveys

examples of high values of ρ^*_{int}, the majority of values cluster below 0.01. The mean values of ρ^*_{int} and the ranges are given in Table 1.

Cumulation of estimates of interviewer effects in figures such as Figure 1, together with other summary values of ρ^*_{int} over surveys and statistics, can demonstrate likely magnitudes of ρ^*_{int} to be expected, given a set of interviewer procedures. Further, as estimates of individual statistics are replicated, similar figures can be constructed for subsets of statistics (e.g., figures for questions on particular domains of attitudes; open questions versus closed questions).

The overall mean ρ^*_{int} is .009 for the nine surveys and approximately 300 statistics examined. Viewing this mean as an average of 300 independent statistics would imply a standard error about it of *approximately* .001, using the guidance of Table 2.[10] Such an average implies an increase in variance of estimates by roughly 10 percent for each 10 interviews added to the interviewers' workloads (relative to the hiring of other interviewers to complete the study). Such results might guide design decisions (following Kish, 1962) to minimize variance of statistics, given fixed resources. The optimal workload is approximately $\sqrt{Ca(1-\rho^*_{int})/c\rho^*_{int}}$, where Ca is the cost of hiring and training an interviewer and c is the cost of completing one interview. For SRC RDD surveys with 20–30-minute questionnaires, Ca is about \$309 and c, about \$18; thus, a ρ^*_{int} of .01 implies an optimal workload of 41 or 42 interviews.

---

[10] This is based on a weighted average sampling variance of .0004575 on individual ρ^*_{int}'s and an assumption of independence among the ρ^*_{int}'s. The lack of independence of the various ρ^*_{int}'s stems from the use of the same interviewers and respondent pairs on multiple statistics of the same survey. It is expected that the correlation among such ρ^*_{int}'s would act to increase the standard error of the mean ρ^*_{int}.

Explaining the Magnitude of Interviewer Effects

In addition to lack of concern about stability of estimates, another weakness in the interviewer effects literature is the failure to seek *explanations* for interviewer effects. The explanations for effects might lie within the nature of interviewer training guidelines, of the interviewer task (e.g., probing instructions, mode of interviewing), or of the survey instrument (e.g., question type). We were able to examine three design features—the form of the question, the type of interviewer behavior used, and the use of computer assistance in the data collection. In addition to these design features we also monitored certain aspects of interviewer behavior and examined characteristics of both the interviewers and the respondents as correlates of interviewer effects.

THE INFLUENCE OF QUESTION FORM ON INTERVIEWER EFFECTS

Previous personal interview surveys have examined question type in an effort to identify the correlates of interviewer variability. Different question types may present to the interviewer communication tasks that vary in complexity. For example, factual questions (those for which there is a knowable, verifiable answer) might be thought to be less subject to effects of different wording or inflection in their delivery than would be attitudinal questions (e.g., "How many times in the last two weeks did you visit a physician's office?" versus "How well do you like TV game shows?"). The results from past studies are mixed, some of them finding attitudinal questions subject to higher interviewer variance. Kish (1962), for example, found no important differences in interviewer effects by question type. However, several studies (Hansen et al., 1961; Fellegi, 1964; O'Muircheartaigh, 1976; Feather, 1973; Collins and Butcher, 1982) have concluded that factual items are less susceptible to interviewer effects. Larger effects have been observed for attitudinal questions, especially those with open-ended responses (O'Muircheartaigh, 1976), for emotionally charged questions (Fellegi, 1964), for difficult items such as income and occupation (Hansen et al., 1961), and for questions which lacked specification regarding the acceptable interviewing procedure (Feather, 1973).

All 297 statistics from the nine different surveys were classified as factual or attitudinal (155 factual questions and 142 attitudinal questions), and since some questions were used in several surveys, there was an opportunity for replicated comparisons. When we examine these results, there appears to be no evidence that factual items as a class are subject to any different interviewer effects than are attitudinal questions. Over all surveys, the average value of ρ^*_{int} was .0098 for factual items compared to .0085 for attitudinal items.

Another variation in question form concerns whether the respondent is supplied answer categories for his choice or is free to formulate his own answer—closed questions versus open questions. Again using data from eight of the nine different surveys, we separated open and closed questions and compared ρ^*_{int}'s. Despite suggestions that open questions might be subject to larger interviewer effects, there was no tendency for this to be the case over 63 open and 192 closed questions examined from the eight surveys. In fact, five of the eight surveys produced higher estimates of ρ^*_{int} for closed questions. Over all surveys, the average value of ρ^*_{int} for closed questions was .0082 compared to .0124 for open-ended responses.

There were, however, aspects of interviewer behavior in using open questions that did vary. For example, some open questions ask the respondent to mention several entities that belong to a certain set (e.g., "What do you think are the most important problems facing the country at this time?"); one statistic of interest is the number of different entities that are mentioned. This count is a function both of the respondent's ability to articulate problems and the interviewer's behavior in probing for more mentions. The number of responses to a question appears to be subject to greater interviewer differences than the substantive response category into which the answers are coded. For example, for six questions asked in the Surveys of Consumer Attitudes, the ρ^*_{int}'s for the percentage giving two mentions are higher each month than the ρ^*_{int} values for the modal response category of the first mention. Over all months, the average value of ρ^*_{int} for the percentage giving two mentions was .0162 compared to an average value of .0007 for the modal category of the first mention. The ρ^*_{int} for the substantive category of the second mention was .0120, close to the .0162 for the proportion with second mentions. This suggests the differential behaviors that determine whether a second mention is given also might influence substantive responses on the second mention.

EFFECTS OF INTERVIEWER BEHAVIOR GUIDELINES ON INTERVIEWER VARIABILITY

In a series of experimental surveys (of which Health in America and the study of Health and Television Viewing are two), interviewers were trained to limit their verbal interaction with the respondent to a specified set of probes or feedback, in addition to the survey questions (following Cannell et al., forthcoming). Most of these statements were entered into the questionnaire itself and were varied to fit the response behavior of the respondent. Interviewers were not permitted to engage in any other conversation with the respondent. Two groups were used: one (experimental) which limited interviewers only to feedback written in a special version of the questionnaire, and another (control) which permitted them to use at their discretion only a small number of statements in response to answers given by respondents (e.g., "I see."). Although both restricted interviewer behavior to a set of specified actions, the first technique was hypothesized to improve reporting (i.e., reduce response bias) over the second.

In both surveys the differences between the ρ^*_{int} values for the cases with specified feedback and those using feedback at the discretion of the interviewer were small relative to the standard errors of the estimates. Since one of these studies (Health in America) has unusually low ρ^*_{int}'s overall, one might speculate that both procedures reduce interviewer effects relative to typical SRC training procedures. A tendency was observed for estimates from the half-sample with specified feedback to be smaller. Combining 57 of the variables from both surveys, 58 percent of the estimates obtained from the experimental group were smaller (mean value of − .0024) than those calculated for the control group (mean value of .0142). Such a finding merits attempts at replication over different substantive topics.

EFFECT OF COMPUTER ASSISTED TELEPHONE INTERVIEWING (CATI)

Another experimental study used terminals to display the survey questions, route the interviewer through the questionnaire, accept, and check on the values of answers recorded by the interviewer. A CATI system thus eliminates certain types of interviewer errors. One study, Health in America, randomly assigned cases to the use of a CATI system or to the use of a paper questionnaire (Groves and Mathiowetz, 1984). Each interviewer used both systems at different times throughout the study. The same questionnaire was used on both systems. Out of the 25 variables compared, the CATI cases have lower ρ^*_{int} values on 18 of the variables, with an average value of −.0031 compared to an average value of .0048 for the paper questionnaire. The differences, however, were small relative to the standard errors of the ρ^*_{int}'s. Attempts to replicate these results are needed.

INTERVIEWER CHARACTERISTICS AND INTERVIEWER EFFECTS

Failure to read a question exactly as printed, inability to follow skip patterns correctly, and reading a question too fast, all are believed to contribute to errors in recorded data (Marquis and Cannell, 1969; Lansing et al., 1971). The Health in America survey monitored interviewers during the survey in order to measure these aspects of interviewer behavior. Two types of behavior, question reading and clarity and pace of question delivery, were analyzed in detail. Even with the special emphasis given to training interviewers for this study, interviewers showed significant variation in the proportion of questions read correctly for many of the variables. For 80 percent of the variables there was a significant difference between interviewers in the proportion of questions read well (i.e., correct pace, clear speech). We found, however, that these differences in interviewer delivery were not correlated with the magnitude of interviewer effects for statistics computed on these variables (e.g., the proportion choosing a particular category).

To search further for this relationship, we created scatterplots of the squared deviation of the individual interviewer's means from the overall study mean by scores from the monitoring data. We expected a positive relationship between an interviewer's squared deviations and the proportion of "incorrect" behavior. No such relationship was apparent. In both the Health in America and the Health and Television Viewing Surveys, we also examined other measures of interviewer performance such as their response rate, productivity, and supervisor evaluations. Again, we found no relationship between any one of these variables and interviewer deviations from survey averages. The so-called "better" interviewers on these characteristics did not deviate any more or any less from the overall mean than did the other interviewers.[11]

---

[11] Most telephone data yield distributions of interviewer means that contain few outliers and have relatively smooth patterns about the overall survey mean. The Health in America survey, however, had many measures where one or two interviewers were extreme outliers to the distribution. Over different statistics the identity of the outliers varied. To evaluate the impact of these extreme values we again performed the interviewer variability analysis for five statistics with high values of ρ^*_{int} (mean value of .0075), eliminating one or two outliers in each case. For each variable the new estimate of ρ^*_{int} was much smaller, with a mean value of $-.0004$ over the five variables. In other words, in that study one or two interviewers were responsible for most of the measured variability.

RESPONDENT CHARACTERISTICS AND INTERVIEWER EFFECTS

For the Health and Television Viewing Survey we examined levels of interviewer variability by sex, education, and age of respondents. The predominance of female telephone interviewers and potential differences in their interaction with female and male respondents suggested the possibility of differences in interviewer effects by sex of respondent. We found little evidence that the gender of the respondent is related to interviewer effects. Although the mean value of ρ^*_{int} over all questions was slightly higher for female respondents, further investigation by question type found mixed results. For some question types the estimates of ρ^*_{int} were larger for males; for other question types the estimates were larger for females. In all cases the differences in the estimates were small.

Although past empirical results vary (Cannell et al., 1977; Schuman and Presser, 1977), it is reasonable to argue that poorly educated respondents might be more easily affected by the behavior and status of the interviewer than more highly educated respondents. Lower education groups may seek greater help from the interviewer in answering the questions or may use the inflection of the interviewer's voice as a cue for responses to questions they find difficult. Values of ρ^*_{int}, however, do not appear to be larger for respondents with less than 12 years of education. Even when estimates of ρ^*_{int} are calculated by question type, differences by respondent education are not apparent.

Respondents were also categorized into five groups according to age. We expected larger values of ρ^*_{int} for the oldest age group. There is some evidence that telephone surveys suffer greater nonresponse among older persons (Groves and Kahn, 1979), and we sought evidence that those who do respond might be more subject to influence by the interviewer because of their own suspicion about the nature of the survey, or greater tendency to fatigue during the 30-minute interview on the telephone. Sudman and Bradburn (1973) have noted that older respondents exhibit greater response errors on questions requiring recall of factual material. In fact, the oldest age group tends to have larger ρ^*_{int} values than other age groups. Over all question types, the oldest age group had an average value of ρ^*_{int} equal to .0274 compared to an average value of .0012 for others.[12] As we noted earlier, ρ^*_{int} measures both variation in nonresponse and response errors over interviewers. These higher ρ^*_{int}'s for the elderly may relate to variation across interviewers in nonresponse rates among the elderly.

---

[12] There is also evidence that the 35–49-year-old group experiences higher ρ^*_{int}'s (mean .0099) but there is no theoretical justification for such a finding.

Summary and Speculations

The instability of estimates from any single survey is the major impediment to measuring interviewer variability. Through cumulating results over many surveys, we obtained generalized estimates of interviewer effects. The analysis showed that for SRC telephone surveys, typical ρ^*_{int} values cluster below .01, an average somewhat smaller than those reported from personal interview surveys, which average closer to .02. We argued that such average ρ^*_{int}'s are useful for guiding decisions regarding how many interviewers to use in order to minimize the variance of estimates, given fixed available resources. In addition, the cumulation of estimates over surveys within the organization could be used to provide more specific information regarding what topics or measurement types might be subject to larger or smaller ρ^*_{int}'s and thus require smaller or larger optimal workloads.

We also attempted to study correlates of interviewer effects. Our results supported others in finding that attitudinal questions are themselves not subject to greater interviewer effects than "factual" items. Open questions are not inherently subject to greater interviewer variability, but the number of distinct answers obtained to an open question appears to be sensitive to interviewer differences. These effects do not apparently influence the substance of the *first* mentioned answer, but appear to be the result of variable probing behavior, that affects both the likelihood of the respondent giving a second mention *and* the substance of that response.

Behavior may be able to be altered to reduce interviewer effects, through use of both controlled feedback and probing and of CATI systems that control the flow and questioning sequences of the interview. These are promising results that deserve further research because they offer the designer real hope of reducing this source of survey error.

In contrast, there was an unfortunate lack of success in using common indicators of performance to predict interviewer effects. Data obtained from monitoring interviewers and the traditional indicators of interviewer quality (response rate, productivity, and supervisory evaluations) seemed to be unrelated to the errors measured by between-interviewer variability. This result casts doubt on the relevance of these indicators for survey quality versus their value for measurement of cost efficiency of surveys.

Finally, we pursued a search for respondent level attributes that were related to interviewer effects. Gender and education of respondent were essentially uncorrelated to the effect, but age of the respondent was related. Older respondents were found to exhibit larger susceptibility to interviewer effects than younger respondents. Given past low response rates among the elderly in telephone surveys, this result might reflect more differential nonresponse bias across interviewers than differential response errors.

Readers of this article need to be cautious in their inference from these results. The studies examined are chiefly health and economic attitude surveys; they use interviewers hired and trained using SRC procedures; supervision, survey administration, and response rates are those achieved by SRC. Since ρ^*_{int}'s can be affected by all these attributes of survey design and practice, these results are applicable to other centralized telephone surveys to the extent their features resemble those of our nine surveys.

References

Cannell, C., L. Oksenberg, and J. Converse
1977 Experiments in Interviewing Techniques: Field Experiments in Health Reporting 1971-1977. National Center for Health Services Research.
Cannell, C.F., R.M. Groves, L.J. Magilavy, N.A. Mathiowetz, and P.V. Miller
Forthcoming "An experimental comparison of telephone and personal health surveys." Vital and Health Statistics, Series 2 Reports.
Collins, M., and B. Butcher
1982 "Interviewer and clustering effects in an attitude survey." Journal of the Market Research Society 25(1):39–58.
Feather, J.
1973 A Study of Interviewer Variance. Department of Social and Preventive Medicine, University of Saskatchewan.
Fellegi, Ivan P.
1964 "Response variance and its estimation." Journal of the American Statistical Association 59:1016–41.
Freeman J., and E.W. Butler
1976 "Some sources of interviewer variance in surveys." Public Opinion Quarterly 40:79–91.
Groves, R.M., and R.L. Kahn
1979 Surveys by Telephone. New York: Academic Press.
Groves, R.M., and N. Mathiowetz
1984 "Computer assisted telephone interviewing: effects on interviewers and respondents." Public Opinion Quarterly 48:356–69.
Hansen, M.H., W.N. Hurwitz, and M.A. Bershad
1961 "Measurement errors in censuses and surveys." Bulletin of the ISI 38(2):351–74.
Kish, L.
1962 "Studies of interviewer variance for attitudinal variables." Journal of the American Statistical Association 57:91–115.
Kish, L., and M.R. Frankel
1972 "Inferences from complex samples." Journal of the Royal Statistical Society (B) 36:1–37.
Lansing, J.B., S.B. Withey, and A.C. Wolfe
1971 Pages 219–318 in Working Papers on Survey Research in Poverty Areas. Survey Research Center, The University of Michigan.
Marquis, K.H., and C.F. Cannell
1969 A Study of Interviewer–Respondent Interaction in the Urban Employment Surveys. Survey Research Center, The University of Michigan.

O'Muircheartaigh, C.A.
1976 "Response errors in an attitudinal sample survey." Quality and Quantity 10:97–115.
O'Muircheartaigh, C.A., and A.M. Marckwardt
1980 "An assessment of the reliability of WFS data." World Fertility Survey Conference, Methodology Session No. 6.
Schuman, H., and S. Presser
1977 "Question wording as an independent variable in survey analysis." Sociological Methods and Research 6(2):151–70.
Sudman, S., and N. Bradburn
1973 "Effects of time and memory factors on response in surveys." Journal of the American Statistical Association 68:805–15.
Tucker, Clyde
1983 "Interviewer effects in telephone surveys." Public Opinion Quarterly 47:84–95.
Waksberg, J.
1978 "Sampling methods for random digit dialing." Journal of the American Statistical Association 73(361):40–46.

The Polls: Environmental Protection

JOHN M. GILLROY AND ROBERT Y. SHAPIRO

SINCE the first Earth Day and the initial peak of the ecology movement in the early 1970s (e.g., Erskine, 1972a, 1972b; Utrup, 1978) the issue of environmental protection has had its ups and downs on the political agenda and in the public's mind. But as a result of recurring problems and serious incidents—Love Canal, Times Beach, PCBs, EDB, dioxin, vinyl chloride, contaminated groundwater, air pollution, acid rain, asbestos, radon gas emissions, the disposal of nuclear/radioactive materials (and problems with nuclear power plants generally), the Superfund and other turmoil at the Environmental Protection Agency, the contamination of wild life preserves, oil spills, and especially the disaster in Bhopal—the environment has been a persistent concern, prominent in the media and magnified by worldwide environmental groups. Hazardous waste disposal and the chemical industry, in particular, with reportedly nearly 7000 accidents involving toxic chemicals in the last five years (Diamond, 1985), are widely considered as serious national problems (e.g., John, 1985). Environmental protection was a leading issue in the last New Jersey gubernatorial election, and this year's national high school debate issue concerned water quality.

Although the Reagan administration has attempted to weaken environmental controls, the American public has been leary of such governmental deregulation. Americans continue to want the government to protect them from economic and personal harm from business and industry (Shapiro and Gillroy, 1984). The evidence (some not presented below because of space limitations) shows a consistently high level of public support for environmental protection, and this is in the shadow of forced budget cuts and the possible weakening of resolve of the Environmental Protection Agency to enforce fully Congress's environmental policy. Congress has reauthorized and invigorated the Clean Water and Clean Air Acts, the Safe Drinking Water Act, the Resource Conservation and Recovery Act, the Ocean Dumping Amendments to the Marine

---

John M. Gillroy is a Congressional Fellow of the American Philosophical Association. Robert Y. Shapiro is Assistant Professor of Political Science at Columbia University and a NORC Research Associate. The authors thank Tom Smith and Sasha Heid for their comments and assistance. Support has been provided by the Columbia University Council for Research in the Social Sciences; the responsibility for analysis and interpretation is the authors'.

Public Opinion Quarterly Vol. 50:270–279 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-270/$2.50

Protection, Research and Sanctuaries Act, and most recently, the Toxic Substances Control Act (with prospects for groundwater and acid rain legislation). In contrast to the executive branch, Congress has perceived and responded to a strong public support for setting and enforcing tough pollution standards.

The survey data during the 1980s reveal substantial and in some cases clearly increasing support for environmental regulation. The year 1983 was pivotal: the Reagan administration had wanted to overhaul the major environmental laws, but it misread or failed to anticipate public opinion (Lake, 1983; Business Week/Harris Poll, 1983; Shabecoff, 1983). In response to survey questions about specific environmental problems, large segments of the public viewed nearly all of them as important *national* issues in need of significant attention, although they may rarely be perceived as *local* problems.

Support for environmental protection has been resistant thus far to energy shortages and price increases, recession, Reagan's landslides, and the high costs of environmental programs and regulation. There was little, if any, appreciable backlash in opinion. The proportion of the public wanting less regulation or less spending has remained small, with relatively little change, even as support for deregulation or "less government" rose at the the beginning of the Reagan administration (cf. Shapiro and Gillroy, 1984). The public has been reluctant to support environmentally harmful policies which might improve the economy and energy resources, with possible exceptions such as increases in strip-mining and offshore oil exploration, although citizens may waver if faced with hard choices close to home (Rosenbaum, 1985:68–70). It is evident, however, that people do not—perhaps rationally—see as mutually exclusive a healthy environment and other important needs, or they do not want to have to choose between them (Ladd, 1982; Mitchell, 1980; Keeter, 1984).

The survey data concerning the costs of environmental regulation and protection are less extensive and revealing (we cannot tell whether respondents will genuinely be willing to pay what they say, although changes in these opinions are suggestive), but this evidence is critical for learning more about the public's policy preferences when the environment's salience wanes after the latest major news event (cf. Downs, 1972). The public almost certainly knows there are significant costs—more so now than ever before—and most respondents say that they would be willing to pay some modest amount (cf. Erskine, 1972a). Of course, the problem of costs has not often been framed in surveys or elsewhere in terms of those paid directly by citizens or by polluters (subsequently paid indirectly by the former). Americans appear to think that those who pollute should pay, and prefer that the public itself contribute through higher prices rather than tax increases (Fiorina, 1973).

One dimension of environmental protection not addressed well in opinion polls is the evaluation of *risks*: the uncertainties of the scientific and technical mechanisms of both the causes of environmental problems and their solutions, and the weighing of specific prospective costs and benefits. What tradeoffs are acceptable? Which must be decided upon soon? Should the government restrict a wide range of activities until all of the long-term as well as short-term risks are assessed? Or are many of these pursuits too important for economic progress, technological advancement, or confronting future problems of scarcity or excessive government spending? How effective are certain government actions likely to be, even with the past as a guide? How much must be done and for how long? These questions may be central to the next generation of environmental issues, and they may have to be grappled with in a public dialogue.

They have evolved from the problems that first burst into the American consciousness more than 15 years ago—and remained.

The survey questions reported here were taken from the files of the Policy Change and Public Opinion Change Project at NORC (originally supported by the National Science Foundation, grant #SES-7912969 A01). These data were compiled from published and unpublished sources, and with the assistance of the Roper Center, University of Connecticut; the Louis Harris Data Center, University of North Carolina at Chapel Hill; the Inter-university Consortium for Political and Social Research, University of Michigan; and the NORC Library. Further information about these data can be obtained from the appropriate archive or survey organization. The national sample sizes are at least 1000 unless noted otherwise.

CRI — Cambridge Reports, Inc., including Cambridge Reports (1985)
HARRIS — Louis Harris and Associates
NBC — National Broadcasting Company
NYT/CBS — *The New York Times*/Columbia Broadcasting System
NORC-GSS — NORC's General Social Surveys
ROPER — Roper Organization
SRC/CPS — Survey Research Center/Center for Political Studies, University of Michigan, American National Election Studies
TRENDEX[1] — Trendex, Inc./General Electric Quarterly Index of Company Relations
YSW — Yankelovich, Skelly, and White

References

Black, Joan S.
1983 "General Electric brings the polls to life." Public Opinion, June/July, p. 27.
Business Week/Harris Poll
1983 "A call for tougher—not weaker—antipollution laws." Business Week, January 24, p. 87.
Cambridge Reports
1985 America Looks Ahead. Cambridge: Cambridge Reports, Inc.
Diamond, Stuart
1985 "U.S. toxic mishaps in chemicals put at 6,928 in 5 years." The New York Times, October 3, p. A1, D22.
Downs, Anthony
1972 "Up and down with ecology—the 'issue-attention cycle.' " The Public Interest 28:38–50.
Erskine, Hazel
1972a "The polls: pollution and its costs." Public Opinion Quarterly 36:120–35.
1972b "The polls: pollution and industry." Public Opinion Quarterly 36:263–80.
Fiorina, Morris
1973 "Public opinion and pollution: an interpretive survey." Environmental Quality Laboratory, California Institute of Technology.
John, Kenneth E.
1985 "Other voices." The Washington Post National Weekly Edition, April 15, p. 38.

---

[1] These surveys were described very briefly by Black (1983), and trend data from them were reported recently in Shapiro and Smith (1985), regrettably without further comment. Unlike most telephone surveys that are done today by the major polling organizations, which use random digit dialing (RDD) methods, the TRENDEX studies interviewed random samples taken from names listed in local telephone directories. Consequently there may be difficulties in drawing inferences from these data about mass opinion, although evaluating *trends* may pose less of a problem.

Keeter, Scott
1984 "Problematic pollution polls: validity in the measurement of public opinion on environmental issues." Political Methodology 10:267–91.
Ladd, Everett Carll
1982 "Clearing the air: public opinion and public policy on the environment." Public Opinion, February/March, pp. 16–20.
Lake, Laura M.
1983 "The environmental mandate: activists and the electorate. Political Science Quarterly 98:215–33.
Mitchell, Robert C.
1980 Public Opinion on Environmental Issues: Results of a National Public Opinion Survey. Washington, D.C.: Resources for the Future.
Rosenbaum, Walter A.
1985 Environmental Politics and Policy. Washington, D.C.: Congressional Quarterly Press.
Shabecoff, Philip
1983 "Ruckelshaus says administration misread mandate on environment." The New York Times, July 27, p. 1, 9.
Shapiro, Robert Y., and John M. Gillroy
1984 "The polls: regulation — part I," and "The polls: regulation — part II." Public Opinion Quarterly 48:531–42, 666–77.
Shapiro, Robert Y., and Tom W. Smith
1985 "The polls: Social Security." Public Opinion Quarterly 49:561–72.
Utrup, Kathryn
1978 "Environmental public opinion: trends and tradeoffs, 1969–78." Resources for the Future, Washington, D.C.

Importance of Environmental Protection

NORC–GSS/ROPER: We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. First . . . Are we spending too much, too little, or about the right amount on . . . Improving and protecting the environment?

NORC–GSS:

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1980*	*1982*	*1983*	*1984*	*1985*
Too little	61%	59%	53%	55%	48%	52%	48%	50%	54%	59%	56%
About right	26	27	31	31	35	33	31	32	31	33	32
Too much	7	8	10	9	11	10	15	12	8	4	08
Dk	6	7	6	5	7	5	6	6	6	4	04
										N=485	*N*=746

ROPER:

	1971	*1973*	*1974*	*1975*	*1976*	*1977*	*1978*	*1979*	*1981*	*1982*	*1983*
Too little	56%	45%	49%	48%	50%	49%	44%	45%	48%	47%	48%
About right	25	30	30	30	31	31	35	35	36	37	38
Too much	5	15	11	13	12	12	13	14	10	10	7
Dk	14	9	9	9	7	8	7	6	7	6	6

ROPER: There are also differing opinions about how far we've gone with environmental protection laws and regulations. At the present time, do you think environmental laws and regulations have gone too far, or not far enough, or have struck about the right balance?

	1973	*1974*	*1975*	*1976*	*1977*	*1979*	*1980*	*1981*	*1982*	*1983*
Too far	13%	17%	20%	15%	20%	24%	25%	21%	16%	48%
Right balance	32	44	37	35	39	36	33	38	38	30
Not far enough	34	24	31	32	27	29	33	31	37	48
Dk	21	15	12	18	14	11	10	10	9	9

CRI: In general, do you think there is too much, too little, or about the right amount of government regulation and involvement in the area of environmental protection?

	1982	*1983*	*1984*	*1985*
Too much	11%	9%	8%	10%
About right	41	34	27	28
Too little	35	44	56	54
Dk	12	14	9	8

TRENDEX: As for their effect on your way of life in the next few years, say within 10 years, how would you rate the importance of each of the following topics . . . very important, of some importance, or not particularly important?

Air pollution control measures

Important	*2/65*	*8/72*	*8/73*	*8/74*	*8/75*	*8/76*	*11/76*	*8/77*	*11/77*	*8/78*	*8/79*	*8/80*	*5/81*	*8/81*
Very	64%	80%	71%	61%	62%	61%	65%	63%	62%	65%	57%	60%	54%	62%
Some	27	15	22	27	25	27	24	27	25	26	29	30	32	27
Not	6	5	7	11	12	11	11	10	13	9	14	10	14	11
Dk/na	2	*	*	*	*	*	*	*	*	*	*	*	*	*

Water pollution control measures

Important	*2/65*	*11/65*	*8/72*	*8/73*	*8/74*	*8/75*	*8/76*	*8/77*	*8/78*	*8/79*	*8/80*	*8/81*
Very	74%	81%	90%	83%	75%	72%	70%	71%	73%	67%	72%	75%
Some	19	14	7	14	19	20	23	22	20	26	19	20
Not	4	3	2	3	6	8	7	7	6	7	9	6
Dk/na	2	2	1	*	*	1	*	*	1	1	*	*

Controlling how land is used by developers

Important	*8/74*	*8/75*	*8/76*	*8/77*	*8/78*	*8–79*	*8/80*	*8/81*
Very	75%	67%	68%	63%	65%	61%	63%	68%
Some	15	23	22	25	27	30	29	25
Not	9	9	9	11	6	8	7	6
Dk/na	1	1	1	1	1	2	1	1

* Less than 0.5%

TRENDEX: I would like to get your opinion on several areas of important *government* activities. As I read each one, please tell me if you would like the government to do more or less or do about the same as they have been on . . .

	Cleaning up air and water pollution		*Controlling hazardous waste*		*Protecting the environment*		
	8/78	*8/80*	*8/80*	*8/81*	*2/82*	*5/82*	*8/82*
More	68%	65%	84%	85%	66%	67%	68%
Same	23	26	12	12	25	24	25
Less	8	9	3	3	9	7	6
Dk/na	1	*	1	1	1	1	1

* Less than 0.5%

TRENDEX: Would you like to see *large companies* do more, do less, or do the same as they have been in connection with . . . Cleaning up air and water pollution?

	8/78	*8/79*	*8/80*	*8/81*
Do more	83%	82%	84%	84%
Do same	14	14	13	15
Do less	2	2	2	1
Dk/na	2	1	1	*

* Less than 0.5%

CRI: In recent years there's been a lot of publicity about hazardous waste disposal sites that are now causing problems. In general, do you think most of the hazardous waste disposal sites that industry and government constructed over the last 30 or 40 years are safe or unsafe?

	1983	*1984*	*1985*
Safe	11%	13%	12%
Unsafe	70	64	72
Dk	19	22	17

CRI: Let's consider for a moment the problem of hazardous wastes. Which one of the industries listed do you think is currently creating the biggest hazardous waste problem?

	1979	*1980*	*1981*	*1982*	*1983*	*1984*
Nuclear power	20	22	22	21	25	25
Automobile	8	5	7	3	1	2
Chemical	27	33	32	34	48	43
Pesticides/insectides	8	9	10	7	8	9
Other	29	25	20	21	12	15
Dk	8	6	8	13	7	7

TRENDEX: How much of a problem are hazardous wastes in your community at present—a major problem, getting to be a major problem, or not much of a problem?

	7–8/80	*8/81*
Major problem	6%	5%
Getting to be major problem	20	17
Not much of a problem	73	76
Dk/na	1	2

Environment, Energy, and the Economy

ENERGY

SRC/CPS: Present government regulations with regard to pollution and other environmental problems limit full use of some energy sources. Do you think the *government should relax environmental protection regulations* to increase the use of these energy sources, or should the *government keep environmental protection regulations unchanged* even though this may delay the production of more energy?

	1–2/80	*4/80*	*6–7/80*	*9/80*	*11–12/80*
Regulations unchanged	39%	45%	44%	45%	46%
Relax regulations	35	28	31	30	29
Relax, qualified	15	14	13	15	12
Don't know	11	13	12	10	12
	N=999	*N*=956	*N*=839	*N*=767	

ROPER: Some people say that the progress of this nation depends on an adequate supply of energy, and we have to have it even though it means taking some risks with the environment. Others say the important thing is the environment, and that it is better to risk not having enough energy than to risk spoiling our environment. Are you more on the side of adequate energy or more on the side of protecting the environment?

	9–10/73	*1/74*	*5/74*	*9–10/74*	*2/75*	*9–10/75*	*10/76*	*9/77*	*9/79*	*9–10/80*	*9/81*	*9/82*
Adequate energy	37%	47%	37%	41%	40%	40%	33%	43%	43%	45%	39%	35%
Protect environ.	37	34	42	39	35	39	44	35	38	36	40	46
Neither, no conflict	12	13	14	12	14	13	13	14	10	12	13	14
Dk	14	6	8	8	10	9	11	8	9	7	7	5

NYT/CBS: What do you think is more important—producing energy or protecting the environment?

	7/1977	*7/1979*
Producing energy	30%	41%
Protecting environment	51	43
No opinion	19	16

YSW: In general, do you feel that President Carter's energy program puts too much, not enough, or just the right amount of emphasis on protecting the environment?

	1977	*1979*[a]	*1980*[a]
Too much emphasis	11%	13%	12%
Right emphasis	43	31	29
Too little emphasis	36	38	45
Not sure	9	17	14

[a] Omits "protecting."

ROPER: There has been much talk about the problem of a fuel and energy shortage in this country. Here is a list of steps that have been or could be taken to conserve supplies of energy (card shown respondent).
The first is . . . Is that something you think we should do or should not do?

	Increase offshore exploration for oil reserves under the ocean				*Allow more strip mining to produce more coal supplies*			
	6/74	*6/75*	*6/76*	*6/78*	*6/74*	*6/75*	*6/76*	*6/78*
Should	68%	70%	61%	68%	46%	48%	42%	53%
Should not	18	13	18	17	37	31	34	31
Dk	14	17	22	15	17	21	23	15

	Eliminate pollution control devices on cars since they use so much more gasoline				*Relax pollution control standards so that fuels which don't meet these standards like coal, high sulphur oil, etc. can be used*			
	6/74	*6/75*	*7/76*	*6/78*	*6/74*	*6/75*	*6/76*	*6/78*
Should	45%	48%	42%	41%	39%	45%	33%	38%
Should not	46	39	44	47	50	41	52	50
Dk	9	13	13	12	11	14	15	13

ROPER: Experts say this all means we must find ways both to conserve oil and to develop new energy sources. Here are some suggestions that have been made (card shown respondent). Would you go down that list, and for each one tell me whether it is something you think should or should not be done. First, . . .

	Increase offshore exploration for oil reserves under the ocean				*Allow more strip mining to produce more coal supplies*				*Relax pollution standards so that fuels which don't meet these standards like coal, high sulphur oil, etc. can be used*			
	2/77	*3/79*	*3/80*	*3/81*	*2/77*	*3/79*	*3/80*	*3/81*	*2/77*	*3/79*	*3/80*	*3/81*
Should	79%	80%	78%	77%	57%	50%	57%	55%	35%	39%	36%	34%
Should not	10	12	14	13	26	37	33	31	51	49	53	51
Dk	11	8	7	10	17	13	11	14	13	13	10	16

CRI: Some people oppose offshore drilling for oil and gas because they fear pollution. Other people say the need for oil is so pressing that we should be willing to suffer an occasional oil spill in order to get more oil. If offshore drilling meant that a major spill would occur about once every 5 years, would you favor or oppose more offshore drilling?

	1980	*1983*	*1985*
Favor	64%	50%	49%
Oppose	22	37	37
Dk	14	14	14

NBC: Sometimes the laws designed to protect the environment make it hard to build a needed oil refinery, a new pipeline or some other energy-related facility. Should the situation arise, which do you think is more important: protecting the environment or building a needed refinery or pipeline?

	9/1979	*10/1981*[a]
Protecting environment	40%	49%
Building energy facility	47	39
Not sure	13	12

[a] The second "needed" is omitted.

ECONOMY

CRI: Which of these two statements is closer to your opinion: We must be prepared to sacrifice environmental quality for economic growth. We must sacrifice economic growth in order to preserve and protect the environment.

	1976	*1977*	*1978*	*1979*	*1981*	*1982*	*1983*	*1984*	*1985*
Sacrifice environmental quality	21%	26%	23%	32%	26%	31%	16%	27%	23%
Sacrifice economic growth	38	39	37	37	41	41	42	42	53
Dk	41	35	39	32	33	28	42	31	24

CRI: Which of these two statements is closer to your opinion? There is no relationship between economic growth and the quality of the environment—indeed, we can have more and more goods and services and also a clean world. We cannot have both economic growth and a high level of environmental quality; we must sacrifice one or the other.

	1976	*1978*	*1980*	*1981*	*1982*	*1983*	*1984*	*1985*
Can have growth and a clean world	50%	47%	50%	45%	49%	49%	55%	51%
Cannot have both	23	30	31	35	35	27	28	33
Dk	27	23	19	20	16	23	17	16

TRENDEX: Now let's talk about how the nation's economy might be affected by various kinds of programs. I will read you a list of programs. As I read each one, please tell me if you think it might *help* the economy, *hurt* the economy, or have little or no effect on the economy . . . An all-out attack on air and water pollution

	1–2/72	*11/75*	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*
Help	71%	58%	55%	52%	55%	53%	55%
Hurt	12	23	22	27	23	26	24
Little, no eff.	14	15	19	18	20	18	18
Dk/na	3	3	5	3	3	3	3

CRI: Would you agree or disagree: Economic policy should encourage growth in those economic areas which are less harmful to the environment and use up fewer natural resources.

	1979	*1984*
Agree	74%	81%
Disagree	10	7
Dk	16	12

ROPER: I want to ask you about some proposals that economists say would have a definite effect on controlling inflation. But while they would help in the long run, they would mean sacrifices and hardships in the next year or so. As I read each one would you tell me if you'd be willing to see it done because of its long-run benefit, or not willing to see it done because the short-term effects would be too severe? First . . . Cut back on pollution and safety regulations that raise the cost of products even if it slows progress on pollution and safety

	9–10/74	*10–11/75*	*10–11/78*	*10–11/79*
Willing	39%	39%	37%	41%
Not willing	46	45	50	46
Dk	16	16	13	13

NBC: Sometimes the laws that are designed to protect the environment cause industries to spend more money and raise their prices. If it comes to a question like that, which do you think is more important: protecting the environment, or keeping prices down?

	6/1978	*12/1978*	*10/1981*
Protect environment	60%	57%	52%
Keep prices down	31	37	37
Not sure	9	6	11

Paying Costs of Environmental Protection

TRENDEX: Would you like to see something done in your community or part of the country on the following issues: . . . If "Yes": If such action were to result in higher costs to you through taxes, utility rates or other prices, how much would you be willing to pay for . . . $5, $50, or $500?

Controlling air pollution

	1–2/66	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	81%	80%	73%	77%	80%	81%	87%	84%	72%	65%	69%	67%
No	19	20	26	21	19	19	12	15	28	34	30	32
Dk/na	*	*	*	1	1	*	1	*	1	1	1	1

	1–2/66[a]	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	49%	45%	44%	43%	41%	40%	35%	36%	43%	44%	40%	41%
$50	27	28	25	30	26	30	33	34	28	29	32	33
$500	6	6	6	9	9	9	12	10	7	6	8	7
Nothing	—	8	9	5	10	9	8	9	11	11	8	8
Dk/na	18	13	15	13	14	12	12	11	11	11	12	12
(*N*)	(813)	(835)	(741)	(825)	(951)	(848)	(969)	(911)	(743)	(670)	(711)	(693)

Controlling water pollution

	1–2/66	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
Yes	90%	83%	90%	84%	89%	86%	91%	93%	91%
No	10	17	10	15	11	13	9	7	9
Dk/na	*	1	*	1	1	1	1	*	*

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	81%	84%	83%	79%	81%	77%	78%
No	18	14	16	20	19	22	21
Dk/na	1	2	1	1	*	1	1

	1–2/66[a]	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
$5	45%	48%	46%	45%	39%	45%	43%	37%	38%
$50	28	23	27	25	30	24	25	33	33
$500	7	5	5	7	8	8	11	10	11
Nothing	—	8	6	5	5	8	9	8	8
Dk/na	20	16	16	17	17	16	12	12	10
(*N*)	(924)		(935)	(849)	(939)		(949)		(973)

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	43%	40%	41%	45%	42%	41%	43%
$50	27	28	29	29	33	33	31
$500	6	6	8	5	7	6	7
Nothing	8	11	9	11	8	7	6
Dk/na	16	15	13	11	10	13	13
(*N*)	(840)	(869)	(844)	(813)	(833)	(783)	(803)

Increase national parks and wilderness areas

	1–2/66	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
Yes	73%	70%	70%	69%	68%	72%	73%	78%	75%
No	26	29	30	29	30	27	26	21	23
Dk/na	1	1	*	1	2	2	1	1	1

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
Yes	68%	69%	67%	65%	68%	63%	63%
No	31	29	31	34	31	35	37
Dk/na	1	1	1	1	1	2	1

	1–2/66[a]	*1966*	*1/67*	*7–8/67*	*1–2/68*	*7–8/68*	*2/69*	*1–2/72*	*1–2/73*
$5	45%	48%	47%	46%	43%	45%	45%	39%	39%
$50	26	21	27	23	28	22	25	31	31
$500	6	5	4	6	7	6	8	10	11
Nothing	—	9	6	5	5	9	9	9	8
Dk/na	23	16	16	20	18	19	13	12	11
(*N*	(754)		(729)	(696)	(716)	(841)	(769)	(856)	(808)

	11/75	*10–11/76*	*10–11/77*	*10–11/78*	*10–11/79*	*11/80*	*11/81*
$5	42%	40%	41%	42%	39%	39%	41%
$50	26	30	29	32	33	33	30
$500	9	7	9	7	11	8	10
Nothing	9	8	9	9	8	8	7
Dk/na	15	15	12	9	9	12	12
(*N*)	(704)	(717)	(684)	(668)	(701)	(646)	(647)

[a] Response categories were coded differently.
* Less than 0.5%.

CRI: How much more per month would you personally be willing to pay for all the goods and services you use as a consumer if you knew that as a result of your paying higher prices business and industry would be able to operate in a way that did not harm the environment?

	1984	*1985*
None	39%	32%
$1–5	19	22
$6–10	11	14
$11–25	9	10
$26–50	4	4
Over $50	4	4
Dk	13	15

NYT/CBS: Do you agree or disagree with the following statement: Protecting the environment is so important that requirements and standards cannot be too high, and continuing environmental improvements must be made regardless of cost.

	9/1981	*9/1982 (Voters)*	*4/1983*	*1/1986*
Agree	45%	52%	58%	66%
Disagree	42	41	34	27
No opinion	13	7	8	7

CRI: Some people say that the government should give more consideration to how much new environmental regulations will cost the public. Other people say that, when it comes to protecting the environment, cost should not be a factor. Which of these views is closer to your opinion?

	1983	*1985*
Consider cost	44%	43%
Don't consider	39	43
Dk	16	13

AOC10-1

Public Opinion Quarterly

Lyndon Johnson and the Public Polls *Bruce E. Altschuler*

Crusading Journalism: Changing Public Attitudes and Policy-Making Agendas *Donna R. Leff, David L. Protess, and Stephen C. Brooks*

Reinterpreting the Gender Gap *Daniel Wirls*

Do Exit Polls Influence Voting Behavior? *Seymour Sudman*

Television's Impact on High School Achievement *Gary D. Gaddy*

Education and the Quality of Life in Korea and the United States: A Cross-Cultural Perspective *Doh C. Shinn*

Types of Political Attitude Structure: Results of a Cluster Analysis *John A. Fleishman*

Economic Partisan Advantages in Congressional Contests: 1938–1978 *Glenn R. Parker*

Response Effects in the Electronic Survey *Sara Kiesler and Lee S. Sproull*

The Polls: Medical Care in the United States *Robert Y. Shapiro and John T. Young*

Proceedings of the Forty-First Annual Conference of the American Association for Public Opinion Research

NEWS AND NOTES • BOOK REVIEWS

FALL 1986

Published for The American Association for Public Opinion Research
by The University of Chicago Press

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

LEO BOGART BURNS W. ROPER BERNARD ROSHCO

ELEANOR SINGER, EX OFFICIO

HERBERT ABELSON AND NORMAN M. BRADBURN, COUNCILLORS-AT-LARGE

Editor: ELEANOR SINGER *Managing Editor:* INA C. JAEGER

Department Editors:

PHILIP A. HARDING, *News and Notes*

MARC B. GLASSMAN, *Book Reviews*

TOM W. SMITH, *The Polls*

EDITORIAL BOARD

JAMES R. BENIGER
JOAN S. BLACK
NORMAN M. BRADBURN
CHARLES F. CANNELL
MURIEL G. CANTOR
ALBERT H. CANTRIL
DONALD DILLMAN
MARTIN R. FRANKEL
HERBERT J. GANS
ALBERT E. GOLLIN
ROBERT HORNIK
MARY R. JACKMAN
F. THOMAS JUSTER
RONALD C. KESSLER
MARK R. LEVY
PHILIP MEYER
J. RONALD MILAVSKY
WARREN J. MITOFSKY
ELISABETH NOELLE-NEUMANN
STANLEY PRESSER
DONALD F. ROBERTS
MORRIS ROSENBERG
CHARLES R. WRIGHT
DANIEL YANKELOVICH

Public Opinion Quarterly (ISSN 0033-362X) Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage. Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540

Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 FALL 1986 NUMBER 3

Lyndon Johnson and the Public Polls *Bruce E. Altschuler* 285

Crusading Journalism: Changing Public Attitudes and Policy-Making Agendas
Donna R. Leff, David L. Protess, and Stephen C. Brooks 300

Reinterpreting the Gender Gap *Daniel Wirls* 316

Do Exit Polls Influence Voting Behavior? *Seymour Sudman* 331

Television's Impact on High School Achievement
Gary D. Gaddy 340

Education and the Quality of Life in Korea and the United States: A Cross-Cultural Perspective *Doh C. Shinn* 360

Types of Political Attitude Structure: Results of a Cluster Analysis *John A. Fleishman* 371

Economic Partisan Advantages in Congressional Contests: 1938–1978 *Glenn R. Parker* 387

Response Effects in the Electronic Survey
Sara Kiesler and Lee S. Sproull 402

NEWS AND NOTES *Philip A. Harding, Editor* 414

THE POLLS: Medical Care in the United States
Robert Y. Shapiro and John T. Young 418

PROCEEDINGS OF THE FORTY-FIRST ANNUAL AAPOR CONFERENCE 429

297

BOOK REVIEWS *Marc B. Glassman, Editor* 454

Stephen Hess, *The Government/Press Connection: Press Officers and Their Offices* *George Bailey* 454

Sawtooth Software, Inc., *Ci2 System 100, Version 1.1* *Frank M. Howell* 455

◆

NOTICE TO CONTRIBUTORS

As of September 1986, manuscripts to be considered for publication should be sent to: Dr. Howard Schuman, Director, Survey Research Center, Institute for Social Research, P.O. Box 1248, University of Michigan, Ann Arbor, Michigan 48106.

Four copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

PUBLIC OPINION QUARTERLY is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

PUBLIC OPINION QUARTERLY has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

Gallup, G.
1972a The Gallup Poll, Public Opinion 1935–1971. New York: Random House.
1972b The Sophisticated Poll Watcher's Guide. Princeton: Princeton Opinion Press.
Goldman, E.
1969 The Tragedy of Lyndon Johnson. New York: Knopf.
Grossman, M., and M. Kumar
1981 Portraying the President. Baltimore: Johns Hopkins.
Harris, L.
1973 The Anguish of Change. New York: W.W. Norton.
Janeway, M.
1976 Review of Wheeler (1978). Atlantic 238:87.
Kearns, D.
1976 Lyndon Johnson and the American Dream. New York: New American Library.
Kernell, S.
1978 "Explaining presidential popularity." American Political Science Review 72:506–22
Mueller, J.
1970 War, Presidents and Public Opinion. New York: Wiley.
Neustadt, R.
1976 Presidential Power: The Politics of Leadership, with Reflections on Johnson and Nixon. New York: Wiley.
Paletz, D.
1976 Review of Wheeler (1978). Political Science Quarterly 91:712–14.
Polsby, N., and A. Wildavsky
1984 Presidential Elections, 6th edition. New York: Scribner.
Roll, C., and A. Cantril
1980 Polls: Their Use and Misuse in Politics. Cabin John, MD: Seven Locks Press.
Sapiro, V.
1977 Review of Wheeler (1978). Journal of Politics 39:1165.
Stimson, J.
1976 "Public support for American presidents: a cyclical model." Public Opinion Quarterly 40:1–21.
Sudman, S.
1982 "The presidents and the polls." Public Opinion Quarterly 46:302–10.
Wheeler, M.
1978 Lies, Damn Lies and Statistics. New York: Dell.

Crusading Journalism: Changing Public Attitudes and Policy-Making Agendas

DONNA R. LEFF, DAVID L. PROTESS
AND STEPHEN C. BROOKS

INVESTIGATIVE reporting is a costly and trendy enterprise used by publishers and broadcasters to build their audiences. Researchers studying mass media influence on the salience of issues before the public—their "agenda-setting" function—seldom have studied this controversial form of reporting. Nonetheless, agenda-setting research is growing, and with it the question of the strength of the causal relationship between media content and public opinion. The methodologies and the findings vary from study to study, with the agenda-setting effects demonstrated both in broadcast and print media. (See literature review by Roberts and Bachen, 1981.)

Abstract This is the third in a series of studies on the agenda-setting effects of news media investigations of social issues. Researchers had access to a local five-part television investigation of repeatedly brutal police officers before it was broadcast in the February 1983 sweeps period. Using random-digit-dialing to select a sample of Chicago area viewers, researchers surveyed 428 members of the general public on their attitudes toward various social issues, including the conduct of city police officers. We hypothesized that respondents exposed to the series would change their responses to questions about police brutality, which they did in statistically significant numbers. A similar survey of so-called elites, policy makers or persons with a prior interest in police behavior, also was undertaken before and repeated after the series was broadcast. Although no statistically significant change in elite opinions of police behavior occurred, small but statistically significant changes in knowledge of specific police acts were recorded. Researchers discuss the findings of study three and speculate about the different agenda-setting effects of the three series.

Donna R. Leff and David L. Protess are Associate Professors of Journalism and Research Faculty, Center for Urban Affairs and Policy Research at Northwestern University. Stephen C. Brooks is Assistant Professor of Political Science at the University of Akron. This work is part of an ongoing project of the Media Cluster at the Center for Urban Affairs and Policy Research at Northwestern, whose members include Fay Cook, Margaret T. Gordon, Paul Lavrakas, Peter Miller and Tom Tyler. The authors gratefully acknowledge assistance from Nancy Trojan and Alan R. Leff, M.D., in manuscript preparation.

Public Opinion Quarterly Vol. 50:300–315 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press
0033-362X/86/0050-300/$2.50

Researchers also have done little to assess the influence of the media on the policy agendas of decision makers (Downs, 1972; Molotch, 1974; and Blanchard, 1974). Studies specifying the connections between the media, public opinion, and policy information tend to be theoretical (Molotch et al., 1982), normative (Paletz and Entman, 1981; Regier, 1930), or anecdotal (Bernstein and Woodward, 1974; Crouse, 1973; Salisbury, 1980).

This study is the third in a series to identify agenda-setting effects of investigative reporting on both the general public and policy makers.[1] All three studies use a pretest, posttest experimental or quasi-experimental design that takes advantage of researchers' prior knowledge of the content of specific investigative reports.

The first in the series (Cook et al., 1983) found that a nationally televised investigative news report on fraud and abuse in the federally funded home health care program had significant effects on both kinds of agendas. The study found that home health-care-related issues (and not unrelated issues) became significantly more important to citizens and policy makers exposed to the television investigative report than to nonviewers. Yet actual policy changes after the televised report resulted more from direct pressure for "reform" by the journalists themselves than from demands by the general public or political constituencies.

The second study (Protess et al., 1985) used a quasi-experimental design to measure the agenda-setting effects of a Chicago *Sun-Times* newspaper investigative series disclosing government improprieties in the reporting and handling of cases of rape against Chicago area women. Again, researchers used a design better suited to test causal hypotheses than methodologies used in more traditional agenda-setting research (Cook and Campbell, 1979), including cross-sectional (McCombs and Shaw, 1972; McLeod, et al., 1974; Erbring et al., 1980), or panel study designs (Tipton et al., 1975; Shaw and McCombs, 1977; MacKuen, 1981). While the first study was of a single nationally televised report using home health care as its topic, the second study was of a local newspaper series on the topic of rape. The third study, reported here, also was of a local multipart series, but the topic was police brutality and the medium was broadcast rather than print.

The agenda-setting effects of the rape study on the general public and on policy makers was more limited than the first study, in part because the pretest disclosed an already high level of awareness and concern about crime among all groups. The most striking result was a sharp increase in the number, length, and prominence of stories about rape in the *Sun-Times* itself—that is, the largest measurable effect was on the medium rather than on its audience. Journalists and government officials did not carefully orchestrate a policy response, as in the home-health-care study. However, as in the first study, policy-making effects were highly "symbolic:" acts "dramatic in outline and empty of realistic detail," rather than "substantive" in nature (Edelman, 1964: 9).[2]

---

[1] The studies are being conducted by an interdisciplinary research team at the Center for Urban Affairs and Policy Research, Northwestern University.

The current study, an analysis of the agenda-setting effects of a five-part television investigative series about repeatedly brutal Chicago police officers, found that the reports had a significant impact on the general public who viewed or heard about them. Response of policy-making "elites" surveyed was less conclusive. Significant changes could be detected in the elites' view of public opinion about police brutality, but they did not increase their own concern about police brutality as a social problem. The series did, however, have an educational impact on exposed elites according to their responses to questions about specific information in the televised reports. The police brutality series resulted also in major policy changes within the Chicago Police Department, in contrast to the largely symbolic gestures that followed the home health care and rape reports.

The first study found greater effects while using a purer experimental design than the second. We therefore were eager to employ the latter methodology in an additional study to determine whether methodological differences accounted for differences in agenda-setting effects in the two studies. We also wanted to study more closely the often overlooked policy effects of investigative reporting and to understand more about the linkages between public opinion, policy makers, and policy formation. This article first will examine the attitudinal impact of the television series on police brutality on the general public and on policy elites. Next, we trace the effects of the series on public policy in Chicago, specifically looking at the police department's own response to disclosures about its shortcomings. We discuss the agenda-setting effect of the series on a heated political campaign in the city, and finally, we analyze the different findings of the three studies and try to explain them.

---

[2] Although some symbolic acts may appear substantive, here we have used the term *symbolic* to describe policy changes that either are purely rhetorical or that would have occurred without media stimulus. Thus, when a public official responds to a media expose by convening a legislative hearing or by announcing as news previously approved legislation, we call those acts symbolic. Similarly, a statement promising reforms that never are implemented is termed symbolic. Conversely, "substantive" reforms are actual regulatory, legislative, or administrative changes that occur after an investigative story is presented.

Research Design

The pretest, posttest experimental design is highly appropriate, but not traditionally utilized, in research involving nonlaboratory studies of media agenda setting. One significant obstacle to its application is inherent in the definition of news. Even media personnel are unaware of specific stories they will cover in the future, making pretesting and experimental or quasi-experimental manipulation difficult or impossible.

In this study, however, two factors made it possible for researchers to apply a quasi-experimental research design. The local NBC affiliate, WMAQ-TV (Channel 5), spent six months investigating allegations of police brutality. While the high drama in the broadcast of the allegations may have made the stories seem immediate, in fact the lengthy preparation gave researchers ample time to obtain preseries measurement of public attitudes toward this problem. Further, the policy relevance of most investigative reporting, such as the police brutality series, makes such stories particularly appropriate for studying their effects on both citizen and policy-making agendas.

The second factor in this study was the close cooperation between Channel 5 reporters and our researchers. In the months before the series aired, researchers were informed continually of the developing investigative findings. Questions could be tailored carefully to the content of the television reports, and issues unrelated to its subject matter could be explored as controls.

The series, "Beating Justice," was broadcast on five successive nights beginning on February 7, 1983—a ratings sweepstakes period coinciding with Chicago's mayoral primary election campaign.[3] The reports included dramatic footage of police wagons used as weapons to bounce handcuffed victims during harrowing rides to jail and interviews with victims maimed by police. Reporters reviewed all lawsuits filed against police in federal courts in Chicago in the previous five years and found 107 officers named more than once. Despite procedures developed by the department in response to earlier investigative reports, officers found by the department to have used excessive force were given only minor discipline and no psychiatric counseling. The series included a statistical review of brutality cases against police describing how the department systematically whitewashed brutality charges. Reporters revealed that millions of taxpayer dollars were being spent to compensate victims of police brutality who successfully sued the department. The cameras graphically depicted—although reporters never explicitly stated—that most police brutality victims were nonwhite. Finally, reporters conducted a live interview with the superintendent of police, who by the conclusion of the series had become an issue in the city's mayoral primary that would be held February 22.

---

[3] The Feb. 22, 1983 primary election was a bitterly fought contest between the Democratic party's endorsed candidate, incumbent Mayor Jane Byrne; State's Attorney Richard M. Daley, son of Chicago's late, long-time mayor (and the sentimental favorite of many old-time party regulars); and black Congressman Harold Washington. Washington defeated his two white opponents, in a primary victory normally tantamount to election in Democratic-ruled Chicago, but in this case the victory led to an even more bitter and racially divisive general election campaign in April against a previously little-known Republican.

Respondents surveyed about offending police officers and the department's lax handling of complaints against them included a random sample of the general public and a purposive sample of policy makers. The methods used in drawing the sample and analyzing data are described below.

GENERAL PUBLIC

Four hundred twenty-eight respondents from the Chicago Metropolitan area were contacted from a pool of potential interviewees using random-digit-dialing techniques. Telephone numbers were generated from listings taken from area directories. The last digit of each number was replaced with a random digit to insure that unlisted telephone numbers were included in the sample. The 30 percent who refused to be interviewed were found to be similar to respondents in sex, age, education, and area of residence.[4]

The sample was stratified according to television viewing habits into watchers of Channel 5 (253 respondents), and watchers of other evening newscasts or nonwatchers of any television news (171). One week after the series was broadcast, researchers recontacted the entire sample. As indicated in Table 1, 279 persons agreed to be reinterviewed, and they comprise the public sample in this study. Those respondents reflected attitudes on the pretest about police brutality that were similar to those on the pretest of persons not reinterviewed, although minorities and the less well-educated were less represented in the reinterviewed group.[5]

[4] About 30 percent of persons "eligible" refused to be interviewed. Eligibles were English-speaking Chicago area residents over the age of 18.

[5] Conducting a difference of means *t*-test between those reinterviewed and those not reinterviewed produced a statistically significant difference between the two groups at the .05 level on 7 of 20 items. The seven included two questions on government corruption, the questions on police bribery and traffic enforcement, and the ratings of police, garbage, and fire services.

We found a difference in the two response groups in race and education (consistent with survey researchers' observations on the mobility of panels of interviewees). The reinterviewed group contained a greater proportion of whites to nonwhites and had a higher education level. Analysis of attitudes indicates, however, that the groups tend to differ little in their initial attitudes about police brutality.

Table 1. Quasi-Experimental Design Distribution of Public Respondents

Respondent Groupings	*Pretest*	*Posttest*
Channel 5 watchers/nonwatchers		
Treatment (watchers)	253 (59.7%)	171 (61.3%)
Control (nonwatchers)	171 (40.3%)	108 (38.7%)
Total	424 (100.0%)	279 (100.0%)
Series aware/nonaware		
Treatment (series aware)	NA	163 (58.8%)
Control (not aware)	NA	114 (41.2%)
Total	NA	277 (100.0%)

At the conclusion of the posttest interview, respondents were asked whether they had "seen, read or heard anything about news media investigations of police brutality in Chicago?" This screening question was used to include in the exposed group persons who may not have seen the series under study but who heard about it in some other way, such as in reports picked up by other media, in conversations with friends, or in advertisements promoting the Channel 5 series. This question also corrected for persons who said they were regular viewers of Channel 5 news but who did not see the series on police brutality. Those responding "yes" to the question were considered "treatment" and those who answered "no" were considered "control" group members. Table 1 shows the distribution of respondents in this analysis. It is important to emphasize that respondents in both analyses are from the same overall sample. It is only the quasiexperimental definition of "treatment" and "control" that changes.

In the pre- and posttest interviews, respondents were asked about a variety of topics including police brutality. Researchers deliberately embedded questions about the police among questions about other social issues to avoid sensitizing respondents to the issues. Of the 24 separate items in the questionnaire, 11 were related to law enforcement issues and 6 to police brutality. We hypothesized that change would occur in the "treatment" group on questions about police brutality, while responses to other questions would remain consistent from pre- to posttest. We expected the "control" group responses to remain consistent on all questions.

We asked respondents to rate the importance of brutality as a problem and to assess the quality of the Chicago Police Department. Further, respondents were asked to agree or disagree with statements about police brutality occurring more frequently against nonwhites, about the police department's lax discipline of brutal officers, and about the need for adequate compensation of victims of police brutality. These questions addressed specific points made in the television report. The subject areas of questions could be divided into four categories: (1) measurement of attitudes and agendas, (2) measurement of knowledge (specific facts learned from the reports), (3) measurement of emotions, and (4) measurement of changes in behavior. Thus, we hypothesized that in posttest interviews, treatment group members would consider police brutality a more important issue, would increase their knowledge of police brutality to incorporate facts revealed in the series, and would change their attitudes, feelings, and actions about the problem.

POLICY MAKERS

A purposive sample of "policy elites" was selected for its interest and potential influence on law enforcement policy makers. The 35 persons surveyed included several suburban police chiefs, state legislators, members of the Chicago city council, and members of a citizens advisory police group. Interviews were conducted by telephone and all 35 were reinterviewed after the police brutality series. In previous studies, researchers made no attempt to establish an elite group of nonexposed ("control") respondents. We expected that persons with a great deal of interest in the topic of investigative reports almost certainly would hear of them, even if they failed to read or view the specific stories. Indeed in the rape study, only one elite had neither read nor heard about the *Sun-Times* reports. In this third study, however, about a third of respondents could be grouped as "controls" who were unaware of the Channel 5 reports. Separate analyses were conducted for the exposed and unexposed groups of elites.[6]

The policy makers were asked a series of pretest and posttest questions that were identical to those in the survey of the general public. Additionally, however, they were asked to estimate the public's views on the issues in question. They also were asked a set of questions about their policy-making activity related to police brutality. After the series, researchers tracked policy changes that might be attributable to the Channel 5 series by interviewing policy makers and by content analyzing local media.

Findings

GENERAL PUBLIC

In examining public attitudes before and after the police brutality series, we wanted to learn whether attitudes toward police brutality changed significantly among persons who watched or heard about the reports. We attempted to measure agenda-setting changes using the two different definitions of "treatment" groups as discussed above.

---

[6] It has been suggested that researchers create a sensitivity control group of policy makers in a future study—a group not subject to the pretest, to determine whether pretest interviewing of elites sensitizes them to the subject of the investigative reports.

Table 2 shows the mean responses of the general public to questions that measured attitude changes and knowledge about specific facts contained in the television reports. The data indicate that persons in the two treatment groups significantly increased their views of the importance of police brutality after being exposed to the reports, although they did not change their ranking of brutality relative to other urban problems. In contrast, persons in the two control groups did not change their opinion of police brutality. However, the strength of these findings was weakened by unexplained significant changes on the importance of welfare fraud (see Table 2), which was the only topic on which control group responses changed. We suspected some media coverage of welfare fraud or a related issue must have occurred during the pretest to posttest interval, but we could find no such coverage. We can only speculate that either some event (that we were unable to document) occurred involving welfare fraud or our survey picked up some other "noise," which we are at a loss to explain. Public attitudes in all groups remained the same on other nonseries issues, although, again, there were changes on welfare fraud and on street repair we cannot explain.

A second attitude question asked persons about the importance of police brutality and four other law enforcement problems not related to material in the series. Again, persons in the two exposed groups showed significant changes in attitude toward police brutality in the posttest survey. Persons in the two nonaware groups showed no changes on any of the issues, as expected.

On factual questions related to the specific content of the police brutality series, change in the exposed groups was weaker.[7] Following the series both treatment groups increased their agreement with a statement that the city compensated victims of police brutality, while persons aware of the series increased their agreement with a statement about the lack of police discipline. Neither of the "control" groups changed their views about any of the content issues relevant to the series, although, as discussed above, both treatment and control groups changed their views significantly on the ability of the government to identify perpetrators of welfare fraud. Neither treatment nor control groups changed their opinions about the veracity of a statement

---

[7] It has been suggested that researchers ask some factual questions about issues *not* covered by the television reports to study possible indirect effects of the story, a suggestion that we or other researchers might want to incorporate in future studies.

Table 2. Mean Responses for General Public[a]

	Channel 5 Watchers/Nonwatchers				*Series Aware/Not Aware*			
	Watchers (Treatment) (N = 171)		*Nonwatchers (Control) (N = 108)*		*Aware (Treatment) (N = 163)*		*Not Aware (Control) (N = 114)*	
	Pre	*Post*	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
Importance of problem								
Police brutality	3.09	3.34*	3.42	3.28	3.18	3.36*	3.28	3.27
Violent crime	3.83	3.81	3.82	3.85	3.80	3.80	3.88	3.87
Welfare fraud	3.50	3.53	3.54	3.35*	3.48	3.38*	3.56	3.56
Unemployment	3.91	3.93	3.88	3.87	3.91	3.93	3.89	3.89
Govt. corruption	3.55	3.54	3.65	3.60	3.58	3.56	3.60	3.57
Importance of law enforcement problem								
Police brutality	3.15	3.38*	3.34	3.29	3.22	3.41*	3.23	3.28
Traffic safety laws	3.02	3.10	3.01	3.10	2.95	3.06*	3.09	3.13
Discrimination in department	2.96	3.05	3.14	2.98	3.02	3.07	3.01	3.01
Underreporting of crime	3.59	3.61	3.51	3.51	3.57	3.65	3.55	3.46
Police bribery	3.60	3.59	3.60	3.51	3.59	3.54	3.65	3.63
Fact statements								
Brutality more likely to nonwhites	2.95	2.88	3.07	3.00	3.02	3.01	2.92	2.81
Lack of police discipline	3.59	3.61	2.75	2.93	2.82	3.11*	2.82	2.68
Brutality victims get money	2.03	2.34*	2.09	2.09	2.05	2.28*	2.10	2.20
Able to identify welfare fraud	2.44	2.65*	2.51	2.79	2.49	2.65*	2.46	2.80*
Punishment for govt. corruption	2.15	2.20	2.08	2.09	1.96	2.09	2.27	2.24
Quality of service								
Police department	3.06	2.98	2.09	2.99	3.00	2.96	2.99	3.02
Garbage collection	3.28	3.35	3.15	3.07	3.25	3.29	3.16	3.18
Fire department	3.66	3.66	3.57	3.53	3.63	3.60	3.58	3.63
Street repair	1.83	2.09*	1.85	1.87	1.86	1.98	1.81	2.00
Parks and recreation	2.82	2.79	2.79	2.72	2.71	2.77	2.91	2.76

[a] Higher numbers indicate more positive responses: (4 equals "very important," "agree strongly," and "very adequate").

* p .05 (asterisks indicate significant change from pre- to posttest using a dependent t-test).

that police brutality victims were more likely to be nonwhite, which was only implied and never directly stated in the Channel 5 series.

Finally, on the question of the quality of the Chicago Police Department's services in general, no similar agenda-setting effect could be detected in any group, although one might have expected a change after the repeated brutality emphasized in the television reports.

To summarize the findings for the general public survey, then, we found significant attitude changes among those exposed to the series. Seven of the 12 tests for related items changed significantly in the hypothesized direction. Such changes were strongest on the general issue questions and weaker on statements of fact. Lack of expected change occurred in respondents' evaluation of police services in general. Tests run on items in which no change was hypothesized showed that 61 of 68 failed to change significantly.

POLICY MAKERS

Table 3 reports the mean responses of policy makers to the pre- and posttest surveys. Responses are divided according to aware and unaware groups, as divided above. Responses are given for policy makers'

Table 3. Mean Response for Policy Elite[a]

	Aware (N = 24)		*Not Aware (N = 11)*	
	Pre	*Post*	*Pre*	*Post*
Personal View				
Importance of problem				
Police brutality	3.00	3.17	2.73	3.09
Violent crime	3.67	3.54	3.91	3.91
Welfare fraud	2.41	2.29	2.63	2.91
Unemployment	3.96	3.92	4.00	3.91
Govt. corruption	3.25	3.00	3.36	3.36
Importance of law enforcement problem				
Police brutality	3.04	3.25	2.73	3.18*
Traffic safety laws	1.78	1.91	1.91	2.09
Discrimination in department	2.82	2.45*	2.56	3.11*
Underreporting of crime	3.18	3.27	3.64	3.72
Police bribery	3.16	3.21	3.64	3.72
Fact statements				
Brutality more likely to nonwhites	2.96	3.26*	3.00	3.63*
Lack of police discipline	2.45	2.95*	2.55	2.91
Brutality victims get money	1.65	2.00	1.80	1.70
Able to identify welfare fraud	2.62	2.19	2.64	2.09
Punishment for govt. corruption	2.50	2.33	2.55	2.73
Public View				
Importance of problem				
Policy brutality	2.64	3.36*	2.63	3.37*
Violent crime	4.00	4.00	3.74	3.87
Welfare fraud	3.09	3.27	2.65	2.74
Unemployment	4.00	4.00	3.96	3.96
Govt. corruption	3.27	3.45	2.78	3.09
Importance of law enforcement problem				
Police brutality	2.81	3.00	2.68	3.00
Traffic safety laws	2.09	2.00	1.86	1.90
Discrimination in department	2.70	3.10	2.70	2.80
Underreporting of crime	3.36	3.55	2.52	3.00*
Police bribery	3.36	3.45	2.86	2.86

[a] Higher numbers indicate more positive response (4 equals "very important" and "agree strongly."

* p .05 (asterisks indicate significant change from pre- to posttest using a dependent t-test).

opinions as well as for their perceptions of the public view. These results may have limited implications due to sample size and nonrandom selection.

Researchers expected less significant changes in elite opinions in all groups because policy makers already are familiar with the issue (and were selected for this sample because of their interest in it). Thus, neither group showed a statistically significant change on the question of police brutality as a general urban problem or as a law enforcement problem. Policy makers who said they had not seen or heard about the Channel 5 reports nonetheless demonstrated a statistically significant change in their opinion on the posttest of the importance of police brutality as a law enforcement problem. (Both aware and nonaware groups in the posttest also increased their views of the importance of discrimination in the police department as a law enforcement problem, which we address in the analysis below.)

The series had its most significant impact for elites on their views of fact statements based on the content of the reports and on their opinions of the public's view of police brutality. Aware elites show statistically significant changes on two of the three fact questions about police brutality. The nonaware groups showed significant changes in their response to the question of how important the public deems the problem of police brutality. They did not change their views of the more specific question about the public's view of brutality as a law enforcement problem.

POLICY MAKING

Policy makers were questioned in the pre- and posttest about their perceptions of the amount of time they are, will be, and should be spending on problems related to police brutality. We hypothesized that after viewing the series, they would increase their time commitment to related issues. While both elite aware and nonaware indicated a limited increase in time commitment, their responses do not allow us to draw any conclusions about their time committed to the problem of police brutality.

We also measured policy-making changes by interviewing public officials after the series and by monitoring the media to learn of changes in policy that might be attributable to the Channel 5 report. The most significant change was a general order, issued by the superintendent of police on March 8, 1983, one month after "Beating Justice" was broadcast. The order established an extensive program for identifying and treating police officers exhibiting "patterns of behavior which would warrant concern," including those against whom complaints of excessive force had been filed. Two new management positions were created to implement the new program. These changes, then, were "substantive" (Molotch et al., 1982), in marked contrast to the highly symbolic responses we found in studies two and three.

Discussion of Findings

With the completion of this third agenda-setting study, we begin to try to make sense of the varying impacts of different investigative reports and to treat some of the methodological issues raised. Because we have completed only three studies, we proceed cautiously; nonetheless, some useful observations are possible.

As in the first study of home health care, this third study found clearly discernible effects in the posttest survey of persons exposed to the investigative reports. Study two, using the *Sun-Times* series on rape, found limited changes in public opinion. The most obvious difference in the studies is the medium used. One and three, which found strong attitude changes, were television reports, while study two was a newspaper report. This is consistent with the growing support for the notion that the impact of television is greater than newspapers when the subject is dramatic and the event short-term (McClure and Patterson, 1976; Graber, 1980). "Beating Justice," with its interviews with brutalized victims and disguised voice, shadow-faced former police officers, certainly was powerful drama. It also was high-quality journalism, as evidenced by the recognition it received, winning the 1983 Chicago "Emmy" for outstanding investigative reporting and the Robert F. Kennedy Foundation award for social justice in media reporting, the first television report so honored. Thus, while the difference in medium may not be the only explanation for the difference in impact, it probably is at least a partial factor and worth exploring in future studies.

The issue itself may also be an important explanatory variable in distinguishing the effects of various investigative reports. It appears that the less information and sensitivity the public has to an issue prior to its treatment in the media, the more likely it is to be influenced by subsequent media reports. Pretests on home health care showed the public to be poorly informed and not concerned about alternatives to institutionalizing the elderly. Similarly, the public surveyed had little concern about police brutality either as a general urban problem or as a law enforcement problem. This lack of prior interest correlated to significant changes in the posttest survey responses in aware groups. In contrast, the public already demonstrated much knowledge and concern about violent crime, including rape, at the pretest stage in study two. It is possible that rape as a social problem already had reached a saturation level on the public's agenda. Thus more opportunity existed for media influence in studies one and three than in study two.

Similarly, it may be argued that police brutality and home health care fraud are "high threshold" issues—discrete problems that are remote from most people's knowledge and opinion formation (Lang and Lang, 1981). Two additional factors may well have increased the issue threshold in the current study. First, Channel 5's emphasis on individual police officers who were repeat offenders provided a specific focus and new angle to the old story of police brutality, distinguishing it in the public's mind from the more general (lower threshold) problems of crime and rape. Moreover, the ability of the two insurgent mayoral candidates to make the series into a campaign issue may have mobilized further public concern about the problem.

The results of the third study also may shed some light on a methodological issue raised by the difference in findings between studies one and two. The first study was an experiment in which great care was taken to maximize internal validity. This was done by random assignment of respondents to treatment and control groups and the use of a selection procedure that would insure that the number of "treated" cases—those who watched the television program—would be large enough for statistical analysis. Since the investigative report was aired on a commercial television show that had low ratings, respondents were asked, and in some cases paid, to watch the program or an opposing program, although they were not directed to watch the particular segment in the one-hour show that pertained to home health care. While this procedure strengthened the internal validity of the experiment, it may have threatened its external validity—the ability of the results to be generalized to similar situations. The request to watch and payment may have prompted respondents to be more attentive to the message (Goetz et al., 1984).

In contrast, studies two and three used quasi-experimental designs, in which respondents were not randomly assigned, but were placed in treatment and control groups based on self-reported reading or viewing habits and responses to questions about their awareness of the series. This was done to enhance the external validity of the studies and because of the difficulty in getting respondents to read or to view a five-part series. The various tradeoffs between internal and external validity in constructing research designs are well documented elsewhere (Cook and Campbell, 1979). In terms of the three studies, the second and third may have stronger external validity because they measure conditions as they more naturally exist in society, while sacrificing somewhat their internal validity.

In analyzing study two, researchers speculated that the agenda-setting effect of the first study (and the limited effects of the second) might be in part due to the fact that respondents were asked to watch and thus observed more closely. The attitude changes among the general public reported in the current study, however, may indicate that the quasi-experimental design also may be effective in measuring the impact of investigative reporting. Future studies are needed to understand whether research design is a critical factor or whether differences such as the nature and presentation of the issue, the medium in which it is presented and its timing in relation to other significant public events are more important. Some unique external events affected the policy-making response to "Beating Justice," especially the bitterly contested Chicago mayoral primary. The "Beating Justice" series was aired only two weeks before election day. Charges of police brutality further fueled a debate among the candidates over the police department's professionalism and accountability.

Race was a major issue in the campaign, in which the white state's attorney, son of the late Mayor Richard J. Daley, and black Congressman Harold Washington attempted to unseat Byrne, the Democratic party organization's candidate. After two nights of the Channel 5 report, Washington seized the allegations of brutality, and the fact that brutality fell disproportionately on nonwhite victims, as campaign issues. He called a press conference in front of the police office responsible for internal discipline, obligating the other news media to cover the issue Channel 5 had raised. He vowed to fire the police superintendent. Later, political analysts cited Washington's exploiting of the police issue in the black community as a factor in his successful campaign. There is little question that the previously described policy responses to the "Beating Justice" series were a direct consequence of the black reformer's election.

Additionally, we need to understand responses to investigative reporting as functions of the media's presentation of an issue. The home health care expose focused on a limited number of unscrupulous operators who already were targets of official investigations. The rape series was largely informational in nature, including statistics on the incidence of rape, victim and perpetrator profiles, and geographic data comparing the city to the suburbs. The police brutality stories, however, practically cried out for reform. Villains and victims were portrayed in unambiguous terms. The evidence presented in the series appeared to be irrefutable. In sum, the stylistic form of an investigative report—equivocal or unequivocal—may affect profoundly its degree of impact. As we study additional investigative reports, we will continue to try to understand the link between the form of the media messages, the nature of the issues exposed, and the opinion and policy responses to them.

References

Bernstein, C., and B. Woodward
1974 All the President's Men. New York: Warner Books.
Blanchard, R. O. (comp.)
1974 Congress and the News Media. New York: Hastings House.
Cook, F., et al.
1983 "Media and agenda-setting: effects on the public, interest group leaders, policy makers, and policy." Public Opinion Quarterly 47:16–35.
Cook, T. D., and D. T. Campbell
1979 Quasi-experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally.
Crouse, T.
1973 The Boys on the Bus. New York: Ballantine.
Downs, A.
1972 "Up and down with ecology: the 'issue attention cycle.' " Public Interest 28:38–50.
Edelman, M.
1964 The Symbolic Uses of Politics. Urbana: University of Illinois.
Erbring, L., E. Goldenberg, and A. Miller
1980 "Front-page news and real-world cues: a new look at agenda setting by the media." American Journal of Political Science 24:16–49.
Goetz, E. G., T. R. Tyler, and F. L. Cook
1984 "Promised incentives in media research: a look at data quality, sample representatives, and response rate." Journal of Marketing Research 21:148–54.
Graber, D.
1980 Mass Media and American Politics. Washington, D.C.: Congressional Quarterly Press.
Lang, G., and K. Lang
1981 "Watergate: an exploration of the agenda-building process." In G. C. Wilhoit and H. deBock (eds.), Mass Communication Review Yearbook. Beverly Hills, CA: Sage.
MacKuen, M. B.
1981 "Social communication and the mass policy agenda." In M. B.MacKuen and S. L. Coombs, More than News: Media Power in Public Affairs. Beverly Hills, CA: Sage.
McClure, R., and T. Patterson
1976 "Print vs. network news." Journal of Communication 26:23–28.
McCombs, M., and D. L. Shaw
1972 "The agenda-setting function of mass media." Public Opinion Quarterly 36:176–87.
McLeod, J. M., L. B. Becker, and J. E. Byrnes
1974 "Another look at the agenda-setting function of the press." Communication Research 1:131–66.
Molotch, H., and M. Lester
1974 "News as purposive behavior: on the strategic use of routine events, accidents and scandals." American Sociological Review 39:101–12.
Molotch, H., D. Protess, and M. T. Gordon
1982 "The media–policy connection." Working Paper Series. Evanston, IL: Northwestern University Center for Urban Affairs and Policy Research.

Paletz, D., and R. Entman
1981 Media Power Politics. New York: Free Press.
Protess, D. L., D. R. Leff, S. C. Brooks, and M. T. Gordon
1985 "Uncovering rape: the watchdog press and the limits of agenda-setting." Public Opinion Quarterly 49:19–37.
Regier, C. C.
1930 Era of the Muckrakers. Chapel Hill: University of North Carolina Press.
Roberts, D. F., and C. M. Bachen
1981 "Mass communication effects." Annual Review of Psychology 32:307–56.
Salisbury, H. E.
1980 Without Fear or Favor. New York: Time Books.
Shaw, D. L., and M. E. McCombs
1977 The Emergence of American Political Issues: The Agenda-Setting Function of the Press. St. Paul: West Publishing Co.
Tipton, L., R. Haney, and J. Basehart
1975 "Media agenda-setting in city and state election campaigns." Journalism Quarterly 52:15–22.

Reinterpreting the Gender Gap

DANIEL WIRLS

In the first years of this decade, political disagreements between the sexes appeared to be the newest cleavage in American politics. Men and women expressed divergent responses to opinion surveys and exhibited different tendencies in the voting booth. These differences came to be known as the gender gap. In the presidential election of 1980, women gave significantly more of their vote to the Democratic incumbent, Jimmy Carter, than did men. During the first years of the Reagan administration, women expressed greater disapproval than men in public opinion polls of Reagan's programs and performance. These and other divisions led the media, interest groups, and many political commentators to describe the gender gap as the product of women's more liberal political values and their greater preference for Democratic candidates, relative to men. Women were said to be react-

---

Abstract This study offers an analysis of the recent political differences between women and men that contradicts the conventional description of the "gender gap" in American politics. Beginning with the 1980 elections, the differences between women and men in their opinions on political issues, in voting behavior, and in partisan identification frequently have been portrayed as a result of women's rejection of Reagan and of Republicans and conservative values and their attraction to more liberal political positions and to the Democratic party. This description is misleading and cannot explain the full spectrum of political behavior by the sexes in the 1980s. This analysis argues that the gender gap has been the result of unequal rates of defection by men and women away from liberal values and the Democratic party. Greater movement by men toward conservative positions and the Republican party produced the gender gap. Consequently, the Republicans were never the potential victims but have always been the potential and actual beneficiaries of the gender gap.

Daniel Wirls is a Ph.D. candidate in the Department of Government, Cornell University. The author wishes to thank Benjamin Ginsberg, whose insight and encouragement initiated this study and improved it throughout the project's duration. Kathleen Frankovic graciously provided CBS News polling data. Martin Shefter, Alice Talbot, and the anonymous reviewers provided helpful comments and criticisms that improved the substance and presentation of the argument.

Public Opinion Quarterly Vol. 50:316–330 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press
0033-362X/86/0050-316/$2.50

ing negatively to Reagan, Republicans, and revitalized conservatism.[1] According to this predominant view, the gender gap spelled trouble for Ronald Reagan in 1984 and signaled the potential coalescence of a women's voting bloc as a key constituent of a new Democratic coalition for the 1980s. This interpretation was the decisive factor behind the nomination of Geraldine Ferraro as the Democratic vice-presidential nominee. In 1984, however, the gender gap did not live up to its putative potential. The difference between the electoral choices of women and men narrowed; had only women voted, Reagan still would have won a landslide victory. Despite the unexpected results in 1984, the conventional interpretation has gone largely unquestioned, though references to the gender gap have become fewer and muted. What explains this contrast between the expectations generated by the popular interpretation of the gender gap in the early 1980s and the reality of 1984?

Proponents of the conventional argument often drew conclusions about the dynamics and future of sex-linked political differences from static measurements of opinion and voting, which showed women to be more liberal or Democratic than men at discrete points in time over the last several years. The history of the gender gap demonstrates the danger of projections based on static data and a conceptual focus on women as the relevant political actors. Longitudinal analysis of the direction of political motion by both women and men produces a much different portrait of the meaning and significance of the gender gap. This study argues that the gender gap resulted from a general movement by the electorate toward Reagan, the Republicans, and more conservative values. While both men and women have been defecting from the Democratic party and moving away from liberal values, rates of defection among men have been greater than rates among women. Hence the gender gap. Rather than reflecting an increase in Democratic and liberal sentiments among women, the gender gap resulted from more rapid and widespread movement among men than women to conservative values and the Republican party. Consequently, Reagan and the Republicans were never the potential victims, but the potential and actual beneficiaries of the gender gap in 1984.

---

[1] For example, Lou Harris said that "consistently, women are voting more Democratic than Republican, but most particularly, are inclined to vote against Ronald Reagan in his bid for reelection" (Ms., July 1984:53). Reagan's chief political advisor during the campaign of 1984, Ed Rollins, warned that the gender gap could "lock" the Republicans "into the status of a minority party," and that it threatened the chances for Reagan's reelection (Washington Post, June 6, 1983). For the perspective of the women's movement see Abzug (1984) and Smeal (1984).

The Data

This study uses three major sources of data. For analysis of electoral behavior the National Election Studies/Center for Political Studies American National Elections Studies from 1976, 1978, 1980, 1982, and 1984 are used. These data were made available through the Inter-University Consortium for Political and Social Research. CBS News/New York Times survey data from 1982 and 1984 on partisan identification were provided by CBS News. For analysis of trends in public opinion, the National Opinion Research Center, General Social Surveys Cumulative File, 1972–1984 is used. This source provides continuity in questions asked, and the compilation of all yearly surveys into a single file facilitates longitudinal analysis.

Interpreting the Gender Gap: The Conventional View

Four specific changes in the political behavior of women and men have contributed to the image of the gender gap and are the basis for its conventional interpretation: the participation gap, the policy opinion gap, the electoral gap, and the partisan gap. Each constitutes a separate gender-based "gap," but together they form the plausible argument for the gender gap as a product of changes in women's political behavior, as a movement away from Reagan and conservatism, if not to liberal values and the Democratic party. The participation gap is, or rather was, the oldest gender gap, for it seems to have disappeared. Until the last decade women voted less frequently and in fewer numbers than did men.[2] The 1970s also marked the closing of the participation gap. In 1980 women voted in higher numbers and, for the first time, in greater percentages than men.[3] The policy opinion gap refers to differences of opinion between women and men on political issues. In addition to the traditional disagreements over the use of military force and violence, polls in the 1970s began to show significant differences across other policy issue areas, such as support for social welfare and environmental protection, with women consistently favoring a more liberal agenda.[4] In 1980 and 1982 these differences over political issues appeared to manifest themselves in a significantly greater preference by women, when compared to men, for Democratic candidates. These election results are the third "gap"—the electoral gap, the original source of the term *gender gap*. In 1980 women gave about 8 percent more of their two-party vote to Carter than did men. In 1982 women voted about 4 to 6 percent more for Democrats and less for Republicans than male voters. Finally, changes in partisan identification provide limited and tentative evidence for a lasting partisan division between the sexes. During the late 1960s and early 1970s, women began to comprise the majority of Democratic identifiers. This margin has increased in the 1980s (Baxter and Lansing, 1980: 67; Frankovic, 1982: 447). The confluence of women's behavior in the polls and in the voting booth gave birth to the heretofore implausible idea of the "women's vote," to be courted by Democrats and Republicans alike. From this perspective, the four "gaps" are intimately related, forming a natural evolution in political consciousness and activity within a changing society. Together they constitute the case for the dominant perception of the gender gap as a product of changing political behavior on the part of women, behavior that supposedly endangers Republicans and favors Democrats. This conclusion was never obvious, however.

---

[2] For discussions of female–male participation differences from the 1950s into the 1970s see Campbell et al. (1960) and Verba and Nie (1972).

[3] For a good summary of census data showing the closing of the gap see Poole and Zeigler (1985: 123–124). The participation gap appears to have been largely an illusion. It was, for the most part, a function of differences in other socioeconomic factors, such as education, age, and income. See Welch (1977), Baxter and Lansing (1983, chap. 2), and Wolfinger and Rosenstone (1980: 42–43).

[4] For analysis of trends in gender-based issue differences, see Baxter and Lansing (1983: 56–60) and Public Opinion (April/May 1982: 21–32).

Reinterpreting the Gender Gap: A Dynamic View

The alternative interpretation is that both women and men are defecting from the Democratic party and liberal values, and the gender gap has been the result of unequal rates of defection, with men outpacing women in an exodus from the liberal and Democratic fold. To test this hypothesis the three principal sources of the conventional interpretation of the gender gap—the policy opinion gap, the electoral gap, and the partisan identification gap—will be examined to show how the data have been misinterpreted and to demonstrate the accuracy of the alternative explanation offered here.

THE POLICY OPINION GAP

While women and men have expressed differences of opinion in polls on various issues of public policy, the differences are not those implied by the popular view of the gender gap. The conventional view of the gender gap emphasizes women's more liberal policy positions. These positions are said to be the source of gaps across several policy areas, specifically the use of force (militarism, defense spending), compassion (governmental involvement in social welfare, health, and promoting equality), and risk (environmental protection, nuclear power). In general, however, this argument was built from measurements of public opinion that revealed the relative positions of women and men to each other while obscuring the overall direction of opinion. The conventional interpretation of gender gaps on political issues is the first example of trends projected from largely static data.

Figures 1 and 2 reproduce data from the NORC General Social Survey from 1974 to 1984. In addition to a general measurement of political ideology, three issue surveys span the policy areas mentioned above. The trend that emerges from the data consistently contradicts the conventional view of the gender gap.[5] Both sexes have moved toward the conservative end of the spectrum. The general political ideology survey summarizes this shift: significantly fewer men and women consider themselves liberals and many more have come to identify themselves as conservatives (see Figure 1). It should be noted that the largest change took place among men, who went from an equal division between liberals and conservatives in 1974 to a 17 percent margin in favor of conservatives in 1984. The specific issue surveys also demonstrate the swing to conservative policy positions and indicate that women and men have not diverged as much as paralleled each other in their opinions (see Figure 2). Across all areas—force, compassion, and risk—both men and women moved away from more liberal values of the early 1970s.[6] A negative reaction to some elements of Reaganism in 1982 and 1983 partially reversed the trend toward conservatism. In most cases the reaction did not return women and men to their degree of support for liberal positions expressed in the 1970s. These reversals in opinion, in fact, only add to the strength of the second point—that the gender gaps in most cases have been stable or have shown no particular trend. Men and women generally parallel each other during cycles of left or rightward swings in national opinion, while frequently displaying a consistent difference of opinion. Even in the supposedly recent divisions over the environment and social spending, NORC data show no consistently growing gaps. In those cases where women and men did swing back toward more liberal positions in 1983, they did so together, and in several cases the gender gaps narrowed.

THE ELECTORAL GAP—THE 1980 ELECTION: WHO MOVED WHERE

The 1980 elections spawned the gender gap. Therefore, a critique of the conventional interpretation must focus on that event. The election

---

[5] This section, admittedly, does not attempt a comprehensive longitudinal survey of sex-linked issue differences. Proof of this point requires further research. What makes such a comprehensive survey difficult to do is a lack of data on a consistent set of policy areas polled over a long time period. Issues come and go and, often, so do the questions. While NORC offers a consistent set of questions, the range of policy issues is limited. Baxter and Lansing (1983) offer a more detailed sketch. Klein (1984: chap. 9) and Poole and Zeigler (1985: chap. 3) also discuss issue differences, but without a longitudinal perspective adequate for addressing the problem in this analysis. Smith (1984) provides longitudinal data on gender attitudes toward violence, war, capital punishment. Although she focuses on change in women's attitudes, Gurin (1985) argues that both sexes have developed more liberal male-female role and sexual equality orientations.

[6] Other issues covered by the NORC surveys such as gun control, capital punishment, governmental spending on health, and race issues show the same trend.

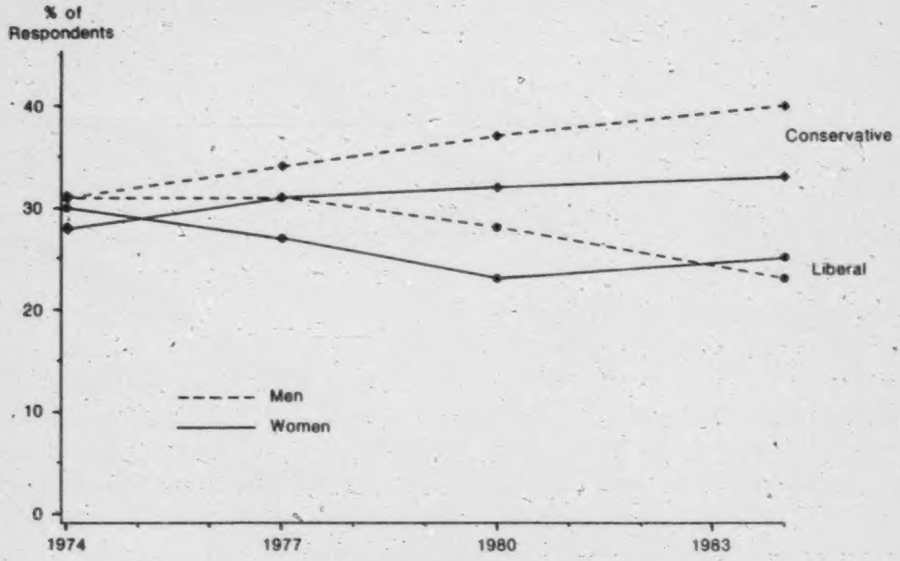


Figure 1. Consider Self Liberal or Conservative

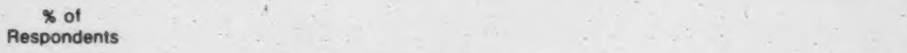


Table 1. The Presidential Vote: 1976, 1980 and 1984[a]

	1976		*1980*		*1984*	
	D	*R*	*D*	*R*	*D*	*R*
Total	51%	49%	44%	56%	42%	58%
Women	51	49	47	53	45	55
Men	51	49	39	61	37	63

[a] Center for Political Studies, American National Election Studies, 1976, 1980, and 1984.

results listed in Table 1 reveal the 8 percent gap mentioned earlier, but the data also uncover a fact frequently ignored in discussions of the gender gap. Women may have voted more for Carter and less for Reagan than did men, but they still gave the majority of their two-party vote (53 percent) to Reagan. In 1976, on the other hand, both sexes gave 49 percent of their vote to President Ford. There was no gender gap in 1976, despite the supposed growth of attitude differences and the female majority among Democratic partisans.

More important, women did move toward the Republicans in 1980, albeit less than did men. Male voters made the dramatic shift in voting behavior, not women. Their vote shifted profoundly—from a slim majority for Carter and the Democrats in 1976 to a 61 percent majority for Reagan in 1980. The female vote can be seen as unexpected, or as a change, only if one anticipated that an incumbent Democratic president would be defeated in a landslide, a very rare occurrence. If men had voted in the same proportions as women, Reagan would still have won, but by a narrow margin reminiscent of other recent elections, such as those in 1960, 1968, and 1976. In other words, the male vote gave Reagan his "mandate" and prompted speculation about the long overdue partisan realignment.

A comparison of women's and men's votes in 1976 with how those same voters voted in 1980, represented in Table 2, produces further evidence that the "movement" is mostly by men and that both sexes

Table 2. How Ford and Carter Voters Voted in 1980[a]

	Reagan	*Carter*	*Anderson*
Ford voters			
Women	80%	13%	7%
Men	82	9	9
Carter voters			
Women	23	67	10
Men	32	62	6

[a] Center for Political Studies, American National Election Study, 1980. CBS News/New York Times data follow the same pattern.

shifted toward the Republicans. Although women may have switched from Ford to Carter in slightly greater numbers than their male counterparts, the two sexes displayed roughly equal loyalty to the Republican party in 1980. There was no significant abandonment of the party by women who voted Republican in 1976. There was, however, a significant divergence between male and female Carter voters, but only in the degree to which they defected from the Democratic party. Men deserted Carter and came over to Reagan in greater numbers than women (by a margin of 9.5 percent). This relative difference between women and men can in no way be construed as a female preference for Democrats. Any party that loses over 30 percent of its former female voters to rival candidates cannot claim to be a lightning rod for the women's vote.

When the 1980 vote is assessed by party identification, the character of Reagan's victory and the gender gap become clearer. Table 3 lists the vote in 1976 and 1980 broken down by sex and party identification. The first striking feature of these data is the similarity in voting between men and women within each of the two major party categories. Democratic women and men divided their votes in almost exactly the same proportions in both 1976 and 1980. Republican men and women did the same. The two-thirds of the electorate comprised by Republicans and Democrats apparently produced relatively little of the gender gap. As we shall see, the independents account for much of the gender gap. The second observation is that the votes of Republican men and women changed very little. The only significant shift appears to have been defections from Carter to Anderson by Republican men and women, with some female defections benefiting Reagan as well. These same changes are apparent among Democratic men and women as

Table 3. Party Identification and the Presidential Vote: 1976, 1980, 1984[a]

	1976		*1980*					*1984*	
	Ford	*Carter*	*Reagan*		*Carter*		*Anderson*	*Reagan*	*Mondale*
Democrats									
Women	18%	82%	22%	[23][b]	72%	[77]	6%	19%	81%
Men	18	82	22	[24]	73	[76]	5	24	76
Republicans									
Women	85	15	89	[95]	5	[5]	5	95	5
Men	90	10	87	[95]	5	[5]	8	96	4
Independents[c]									
Women	54	46	49	[59]	35	[41]	16	67	33
Men	55	45	65.5	[75]	21.5	[25]	13	73	27

[a] Center for Political Studies, National Election Studies, 1976, 1980, and 1984.

[b] Bracketed data are the two-party vote percentages.

[c] Combines the Independent and No Preference categories of the CPS partisan identification question.

well—Reagan and Anderson gained votes as Carter lost support. Once again, the data show that Republican women did not join the Democratic camp and Democratic women did not close ranks for Carter in 1980.

The most significant change represented in Table 3 is within the independent vote. Independents went from a 54.5–45.5 percent margin in favor of Ford, the loser in 1976, to a 58–28 percent margin in favor of 1980s winner, Ronald Reagan. More important is that, unlike the Democratic and Republican men and women, the independents split quite sharply along gender lines. Women independents were about 15 percent less likely to vote for Reagan and 15 percent more likely to vote for Carter than were their male counterparts. The movement by both sexes within this relatively moderate-conservative group (after all, they voted 55–45 percent for Ford) was toward the Republicans. Independent women added to Reagan's victory, but only marginally so, in comparison to the contribution of independent men. Two important facts about Reagan's victory and the gender gap emerge from Table 3. First, the independent vote contributed disproportionately to Reagan's margin of victory. According to the Center for Political Studies data, Reagan won by a margin of about 12 percent (56–44 percent). When independents are excluded from the data, Reagan's margin of victory shrinks to about 4 percent (52–48 percent). Only about one-third of Reagan's margin of victory resulted from Democratic desertion and increased loyalty among Republicans. Reagan may have owed two-thirds of his "mandate" to the independents and otherwise uncommitted. Secondly, the largest gender gap was between independent women and men. The chief partisan source of Reagan's 1980 landslide was also the locus of the widest gender gap. This apparent anomaly can be explained by rejecting the popular image of the gender gap and embracing the opposite interpretation: the gender-based voting differences in 1980 represented different degrees of deviation on the part of men and women away from their previous support of Democrats. The data show that the vote in 1980 portended no nascent boon for the Democratic party. No Republican need fear, nor can any Democrat take comfort from, a gender gap wherein women give a smaller majority of their vote to the GOP and a bigger minority of their vote to the Democrats than their male counterparts.

THE GENDER GAP IN THE 1982 ELECTIONS

The results of the 1980 presidential elections do not end the story of the gender gap. In the 1982 congressional elections another gender gap occurred. Women voted about 6 percent more for Democrats and 6 percent less for Republicans than did men. But as shown in Table 4,

Table 4. House Vote: 1978 and 1982[a]

	1978		*1982*	
	D	*R*	*D*	*R*
Women	58%	42%	60%	40%
Men	59	41	54	46

[a] Center for Political Studies. National Election Studies, 1978 and 1982.

both sexes gave the majority of their votes to Democrats. This was not unexpected. The House has been Democratic since 1954. Furthermore, the party of the incumbent president often suffers during midterm elections. Therefore, the 1982 elections do not represent a resurgence of Democratic vitality. Advocates of the female-oriented view of the gender gap, however, point to the 1982 elections as being indicative of women's movement to the Democratic party. If women led a shift back to the Democrats in 1982, that would be evidence for the popular interpretation of the gender gap. But the data do not support this. Comparison of results from 1978 and 1982, the last two midterm elections, indicates that the two-party division of women's votes for House candidates changed very little between 1978 and 1982. Men, on the other hand, appear to have drifted away from the Democrats and to the Republicans. While this is not conclusive evidence that the gender gap is predominantly the result of a transitory attraction by men to the GOP because of Reagan, or even a longer term commitment by men to the Republican party, the data certainly do not support the dominant interpretation of the gender gap.

THE 1984 ELECTIONS: WOMEN, MEN, AND THE REAGAN ADMINISTRATION

Some of the strongest evidence for the popular perception of the gender gap came not from actual electoral behavior or from partisan preferences, but from differences in political attitudes as expressed in public opinion polls. These were the differences that fueled Republican fears and Democratic hopes as the elections of 1984 approached. Synoptic of these recent divisions between the sexes was the gender gap over approval ratings of Reagan's performance in office. From 1981 through 1983 the Gallup polls showed an average gap of 8 to 10 percent between men's higher degree of approval and women's lower ratings (see Gallup Report, #219, December 1983). These same polls also showed significant gaps on Reagan's handling of several specific issues, such as foreign policy and the economy. Early in 1984, however, these trends failed to sustain themselves, and as the election season progressed it was clear which party would be on the winning side of the gender gap. Women gave Reagan a 51 percent positive rating in the February 1984 Gallup presidential approval ratings, exceeding the 50 percent mark for the first time since the fall of 1981 (see Gallup Report, #220/#221, January/February 1984). In a Gallup poll taken in early September of 1984, which pitted Reagan and Bush against Mondale and Ferraro, men went for Reagan–Bush by 59 percent to 33 percent, but women also gave the majority of their "votes" to the Republican ticket by 53 to 41 percent (see Gallup Report, #228/#229, August/September 1984).

The 1984 elections made clear the inaccuracies of the conventional interpretation of the gender gap and confirmed the alternative description offered here. Table 1 shows that the gender gap was about 8 percent.[7] In 1984 there was no doubt about the direction of the women's vote. Women went from a small majority in favor of Reagan to a 55 to 44 percent split. Table 3 disaggregates the vote by sex and party identification. Republican women increased their loyalty to Reagan. Democratic men and women each gave a slightly larger percentage of their votes to Mondale than they did to Carter, without diminishing their level of support for Reagan. This apparent increase in Democrats' support for the candidate of their party is not indicative of greater loyalty as much as depleted ranks. Many of those who in 1980 called themselves Democrats but voted for Reagan must have become part of the swelled Republican ranks by 1984, thereby paring down the Democrats to a notably smaller cadre.[8] Independent women swung over to Reagan in large numbers. Once again, the biggest gender gap was within this independent group of voters—a group that nonetheless voted 70 to 30 percent for Reagan overall. The gender gap in 1984 was not a tug of war between the sexes over the presidency. Both women and men were pulling on the same end of the rope.

The diminished gender gap at the presidential level in 1984 led analysts to seek confirmation of the conventional interpretation elsewhere, especially at the state level of congressional and gubernatorial races. Ethel Klein (1985) uses network poll data to show that the women's vote provided the margin of victory for Democrats in at least four

---

[7] CBS News/New York Times reported a gap of 4 percent, the Los Angeles Times: 7 percent, NBC News: 9 percent, and ABC News: 8 percent (see Public Opinion, January 1985: 28–29). It is therefore unclear whether the gender gap shrank. Nevertheless, it did not grow, despite indications that it would. Regardless of the size of the gender gap, the strength of the women's vote for Reagan and Republicans illuminates the weaknesses of the conventional viewpoint.

[8] Identification with the Democrats dropped from over 40 percent in 1980 to under 38 percent in 1984, while Republican support increased in the same period from about 23 to about 29 percent. See Table 5.

Senate races. Women made the difference for some Democratic candidates in close races. To take one example, however, if in fact it took the women's vote in Massachusetts to assure John Kerry, the popular Democratic lieutenant-governor, of victory in perhaps the most Democratic state in the nation—a state that voted for Reagan in 1984—then the women's vote is not a wellspring of support as much as it is a part of an evaporating constituency. Finally, CBS/New York Times reported that men gave only 46 percent and women 48 percent of their votes to Democratic congressional candidates.[9] The decline of women's support for Democratic congressional candidates to below 50 percent and the small size of the gap do not indicate a rejection of Republicans and conservatism, let alone a liberal voting bloc.

THE PARTISAN IDENTIFICATION GAP

Two trends in partisan identification are often cited in support of the conventional interpretation of the gender gap. First, women have come to compose the significant majority of Democratic identifiers. Secondly, in the 1980s young men and women have diverged radically in their preferences for the two major parties. Both of these trends, however, turn out to be quite problematic for the popular view of the gender gap and give further support to the argument offered here. For this analysis the interesting and significant point revealed in the data on partisan preference (reproduced in Table 5) is not that women now compose about 57 percent of all Democratic identifiers, but how they came to make up that growing majority. The data indicate that while women have maintained a stable loyalty to the Democratic party, men have defected in significant percentages. Women have become the majority of the Democratic party by default and not by active realignment.[10]

The second trend revealed in recent polls and noted by Frankovic is among young adults: "In the 1982 aggregated polls, the men between the ages of 18 and 24 divide 31 percent Republican and 24 percent Democratic . . . [whereas] 21 percent of women 18–24 call themselves Republicans, while 33 percent say they are Democrats" (1982: 447). This gap between young men and women is significant and surprising.

---

[9] CBS poll data taken from Public Opinion, January 1985, p. 31. Klein (1985) cites ABC News and NBC News as reporting a gap of 6 percent, with men voting 47 percent and women 53 percent for Democrats. If true, the more important fact would seem to be the decline of men's support for the traditional majority party to below the 50 percent mark.

[10] Gallup polls from 1979 to 1984 confirm this trend and even show significant defection by women from the Democratic party. See Gallup Opinion Index, #168, July 1979 and Gallup Opinion Report, #228/#229, August/September 1984.

Table 5. Party Identification: 1980, 1982, and 1984[a]

	1980		*1982*		*1984*	
	Men	*Women*	*Men*	*Women*	*Men*	*Women*
Republican	23%	22%	25%	26%	29%	28%
Independent	38	36	39	30	36	29
Democrat	39	42	36	44	35	43

[a] Figures for 1980 taken from Frankovic, 1982: 447; all other years aggregated by the author, using CBS News/New York Times polls.

The most provocative fact is not that young women are dividing themselves between the two parties in roughly the same proportions as all adult Americans have for many years (that is, pro-Democratic by margins well over 10 percent). Rather, the astounding fact is that young men are now going against the grain of the American electorate of recent history by identifying themselves as Republicans much more than as Democrats. Once again, the evidence leans toward the male-oriented explanation of the gender gap and away from the female-Democratic interpretation. The argument here is not that women are not becoming an important part of the Democratic constituency. In some ways they are. The important consideration is that the Democrats are not experiencing a net gain in support as a result of the relative changes by the sexes in political outlook and voting behavior. The Democrats are not gaining women as much as they are losing men.

Discussion

By amending the popular misapprehension of the gender gap, this study is intended to contribute to the growing debate on sex and politics. It explains the manner in which the sexes diverged in political behavior. It does not purport to offer a complete explanation for the gender gap. This analysis leaves a crucial stone unturned. The question remains as to why the sexes diverged in the 1980s. This debate is important and has received much scholarly attention.[11] Relatedly, the gender gaps in voting behavior and public opinion are not solely functions of the single variable of sex.[12] More multivariate analysis is needed to uncover the other social and economic forces behind the gender gap.

[11] To cite a few examples: Frankovic (1982), Klein (1984: chap. 9), and Mansbridge (1985). The focus of analysis here is on the linkage between gender-based issue differences and voting behavior.

[12] In addition to those cited above, see Monk-Turner (1982), Blydenburgh and Sigel (1983), Goertzel (1983) and Burris (1984). For example, Bolce (1985), whose thesis is similar to the one in this study, examines the impact of race on the gender gap.

The conventional interpretation of the gender gap sprang from a narrow empirical and conceptual focus to become a central concept in the political vocabulary of the 1980s. The results were unwarranted expectations in anticipation of November 1984. Contemporary polling technology quantified, and then the media publicized, the substantial difference between the voting behavior of women and men in 1980. All too quickly the language used to describe this phenomenon converged around a dominant interpretation—an interpretation that explained the gender gap as movement by women within the political arena. This interpretation was seized upon by those political actors who benefited from its implications, namely the recently defeated Democrats and women's groups. The Democrats needed the support of new constituencies in order to rebuild a coalition capable of recapturing the presidency. Women's political action organizations used the leverage this issue gave them to increase their influence within the Democratic party and within their constituency. In the short run, the strategy worked well for the women's movement. Its power within the party increased dramatically and was personified in Geraldine Ferraro and her candidacy. But the strategy was an electoral failure for the Democratic party. The election of 1984 exacerbated the tensions within the traditional Democratic coalition and showed that the gender gap was strategy based on illusion. In the long run the gender gap myth may benefit neither group. The women's movement may have overplayed its cards and the Democratic party will almost certainly search for other sectors of support. The use of the gender gap for partisan causes was and is unavoidable. Political opponents should do battle with the weapons provided by social science. Nevertheless, the politicization of social science has its costs, and they are evident in the case of the gender gap. The dominant interpretation of the gender gap misled analysts and politicians and obscured the more accurate and important meaning behind the numbers.

References

Abzug, Bella
1984 Gender Gap. Boston: Houghton Mifflin Company.

Baxter, Sandra, and Marjorie Lansing
1983 Women and Politics. Ann Arbor: University of Michigan Press.

Blydenburgh, John C., and Roberta S. Sigel
1983 "Key factors in the 1982 elections as seen by men and women: an exploration into the vulnerability thesis." Paper delivered at the 1983 Annual Meeting of the American Political Science Association, Chicago.

Bolce, Louis
1985 "The role of gender in recent presidential elections: Reagan and the reverse gender gap." Presidential Studies Quarterly 15: 372–85.

Burris, Val
1984 "The meaning of the gender gap: a comment on Goertzel." Journal of Political and Military Sociology 12:335–43.

Campbell, Angus et al.
1960 The American Voter. New York: Wiley.

Frankovic, Kathleen
1982 "Sex and politics—new alignments, old issues." PS 15:439–48.

Goertzel, Ted G.
1983 "The gender gap: sex, family income, and political opinions in the early 1980s." Journal of Political and Military Sociology 11:209–22.

Gurin, Patricia
1985 "Women's gender consciousness." Public Opinion Quarterly 49:143–63.

Klein, Ethel
1984 Gender Politics. Cambridge, MA: Harvard University Press.
1985 "The gender gap: different issues, different answers." The Brookings Review 3 (Winter):33–37.

Mansbridge, Jane
1985 "Myth and reality: the ERA and the gender gap in the 1980 election." Public Opinion Quarterly 49:164–78.

Monk-Turner, Elizabeth
1982 "Sex and voting behavior in the USA." Journal of Social, Political and Economic Studies 7:369–76.

Nie, Norman, Sidney Verba, and John R. Petrocik
1979 The Changing American Voter. Cambridge, MA: Harvard University Press.

Poole, Keith T., and Harmon L. Zeigler
1985 Women, Public Opinion, and Politics. New York: Longman.

Public Opinion
1982 "Men and women: is realignment under way?" April/May, pp. 21–32.
1985 "Opinion roundup: the 1984 election results" December/January, pp. 23–42.

Smeal, Eleanor
1984 Why and How Women Will Elect the Next President. New York: Harper and Row.

Smith, Tom W.
1984 "The polls: gender and attitudes toward violence." Public Opinion Quarterly 48:384–96.

Steinem, Gloria
1984 "How women live, vote, think." *Ms.*, July, pp. 51–4.

Verba, Sidney, and Norman H. Nie
1972 Participation in America: Political Democracy and Social Equality. New York: Harper and Row.

Welch, Susan
1977 "Women as political animals? A test of some explanations for male–female political participation differences." American Journal of Political Science 21:711–30.

Wolfinger, Raymond E., and Steven J. Rosenstone
1980 Who Votes? New Haven, CT: Yale University Press.

Do Exit Polls Influence Voting Behavior?

SEYMOUR SUDMAN

EXIT POLLS are simply the latest technology used by the news media to predict elections before all the votes are in. Before exit polls, the media concentrated on key precincts, and before that simply on experienced political journalists who made predictions based on comparisons of known voting patterns to those of earlier elections. The type of information reported is not new. What is different is that the current technology is far more accurate and faster, since the data are collected and analyzed while the election is in progress.

Many politicians, political scientists, and communication researchers have tried to demonstrate that election evening news does or does not have an effect on whether people vote and how they vote. In this article, we summarize their findings and discuss the major statistical problems with these studies.

There have been essentially two procedures used to measure media effects on voting, a macro approach that compares total voter turnouts in different areas or elections and a micro approach that uses surveys of individual voters and nonvoters and determines from these respondents whether they had been exposed to election evening news (Dubois, 1983; Epstein and Strom, 1981; Fuchs, 1966; Tannenbaum and Kosterich, 1983). Underlying all of these studies is a model of voting behavior that is worth making explicit, since it will help us to analyse what is right and wrong about the studies that have been done.

Abstract This article summarizes the findings of the effect of exit polls on voting behavior. Both macro and micro methods have been used, and no methods are perfect. Exit polls appear to cause small declines in total voting in areas where the polls close late for those elections where the exit polls predict a clear winner when previously the race had been considered close.

Seymour Sudman is Walter Stellner Chair of Marketing at the University of Illinois at Urbana/Champaign. An earlier version of this article was presented at the Annual Meetings of the American Statistical Association, Las Vegas, Nevada, August 1985.

Public Opinion Quarterly Vol. 50:331–339 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-331/$2.50

Most political scientists see voting as a behavior like other human behaviors, with both benefits and costs. Citizens vote when they perceive the benefits of voting as clearly outweighing the costs. More explicitly, some of the benefits of voting are related to (1) the perceived importance of the election, (2) the perceived closeness of the election, and (3) more general ideas about the importance of voting in democracies. That is to say that there are some people who vote regardless of the characteristics of the election, but many who balance the value of their vote against the perceived costs among which are (1) time required—the distance and convenience of the polling place, and the waiting time required to vote, (2) interference with other activities, and (3) discomfort (getting wet while standing in line).

When deciding whether or not to go to the polls, the potential voter is confronted by elections for a large number of positions ranging from president to senator to representative down to obscure local offices. The final decision to go to the polls is made on the basis of some decision function that combines all of these races.

Where do the exit polls fit into this? Obviously, they can influence voters' perceptions about the closeness of the election and the value of their votes. Think of this extreme scenario:

A Californian plans to vote after work in what she believes to be a close presidential election. (She has little interest in the race for congressman for her district, although it is closer.) The day is rainy and as she approaches the polling place she sees a long line. On the radio she hears that one presidential candidate has a substantial lead in other states. She says why bother and turns her car around and drives home.

Although this is a completely fictional description, it has almost certainly occurred in the past and will occur again in the future. Thus the question is not really do the exit polls have *any* effect on voting behavior, but what is the magnitude of that effect? Are 1 percent of voters affected, 10 percent, or more? The studies that have been done range from showing no effects to showing more than 10 percent of voters potentially affected. Obviously, these cannot all be right, or they may be looking at different elections with different characteristics.

An important issue is the perceived closeness of the election before the election occurred and the peceived closeness of the election based on exit poll results. Jackson (1983), whose paper we discuss in greater detail below, points out very sensibly that the only time the exit polls can have an effect is when they change voters' perceptions about the closeness of the race. This was the case in 1980, when a close race was expected based on preelection polls, but the exit polls showed a Reagan sweep. This was not the case in the 1984 elections, where the exit polls simply confirmed what the preelection polls had already forecasted, a wide Reagan victory. Nor was it the case in the 1976 elections, which had been predicted as close and were also reported by the media as close on Election Day. In fact, the only election like 1980 may have been the 1960 election, where the preelection polls had a tight contest but where the early returns from the East Coast gave President Kennedy a substantial lead. As Jackson points out, the exit polls may have the converse effect of causing more people to vote if they indicate the race is closer than expected. Thus, in 1948 and 1968, where one-sided elections were expected, exit polls could have brought people in to vote by indicating close elections.

The observed relation between turnout and perceived closeness of the election does not, of course, mean that every voter behaves in the same way. Some voters may be encouraged to vote when the election is not close because they are sure to be on the winning side (a bandwagon effect). Others may vote to reduce the magnitude of the anticipated winning margin (an underdog effect). Overall, however, neither bandwagon or underdog effects have been observed in national elections.

In any event, this discussion suggests that one needs to look at the characteristics of elections and that for many elections no effect would be expected. The 1980 election, because of its special features, and because it was the first one in which exit polls were widely used, has been examined most often, and we limit our discussion to this election.

The Micro Level

On a micro level, the paper by Jackson is the only one available for the 1980 election. There had been earlier uses of individual data by Mendelsohn (1966) in an analysis of the 1964 election, but since this had been predicted as a very large victory for President Johnson and the early results indicated the same thing there was no reason to expect an effect, nor was there any. A recent paper by Adams (1985) about voting behavior in Oregon during the 1984 presidential elections also found no effects since the substantial victory for Reagan had been anticipated.

Jackson used data from the University of Michigan Presidential Study in 1980. Respondents were interviewed before the election to determine if they were registered and their likelihood of voting. A sample of 1814 were reinterviewed two months after the election and were asked about their exposure to different kinds of news coverage: The exit polls, Carter's concession speech, media projections of the winner, election results, and newspaper results. The actual voting behavior of respondents was obtained by checking voting registers, since it is well known that some people claim to vote who did not do so.

Given the data, Jackson estimated the change in likelihood of voting for a person who had not voted before 6 PM EST as a linear function of the effect of hearing various aspects of the election coverage. Perhaps the easiest way to summarize his results is to show his table on estimated turnout probabilities by region based on whether respondents had or had not heard election results (Table 1). To read this table, Jackson predicts that 43 percent of registered voters in the West would have voted before 6 PM EST and that after 6 PM 50 percent would have voted if no one heard news of the election, but only 38 percent would have voted if everyone had heard the news.

From this he concludes that up to 12 percent of voters could have been affected by the early news *if they had heard about it*. His results, which unfortunately are not given by region or voting status indicate that only one of six persons interviewed had actually heard any of the exit poll results before the polls closed in their state. If one used this figure for nonvoters everywhere, it would suggest that between 1 and 2 percent of potential voters did not vote because they heard the results of the exit poll. This figure could be as high as 3 or 4 percent in the West if the percentage who had heard the exit polls results was much higher than everywhere else. In a just published study, Wald (1985) presents evidence that nonvoters are less likely to know whether an election is or is not close and thus unlikely to have heard or been affected by exit polls.

There are several problems with the Jackson approach, some of which he recognizes. The first is that it depends on respondents' memory of when they voted and heard or saw news two months afterwards. This is an extremely difficult memory task subject to substantial error. Jackson argues, however, that it would affect his estimates only if nonvoters had worse memories than voters, an event he finds unlikely. Jackson may be right, but it is certainly possible that some social desir-

Table 1. Estimated Turnout Probabilities

	Average Regional Respondent			
	Before 6 PM EST	*After 6 PM EST*		
		Not Heard	*Heard*	*Difference*
East	0.81	0.14	0.08	0.06
Midwest	0.56	0.22	0.10	0.12
South	0.70	0.18	0.09	0.09
West	0.43	0.50	0.38	0.12

ability effects may be causing differential memory errors between voters and nonvoters. Nevertheless, this is not the most serious problem.

A second problem inherent in any micro approach is in the rapid shrinking of sample sizes. Of the initial sample, only 357 respondents did not report voting before 6 PM EST. When this sample is split by region, and data are given by likelihood of voting and having heard exit poll results, the results must be based on very small samples. This does not mean that the results are biased, but that they are unstable. The regional results that show no differences between the Midwest and West may be real, but to me they suggest the effect of small samples.

Most critically, Jackson makes the assumption that the analysis closely resembles a natural experiment, with some subjects receiving the news when it could alter their behavior (the experimental group) and others not being so exposed (the control group). Unfortunately, it is equally likely that the behavior that was to have been affected, voting, was actually the cause of whether or not news was heard. That is, a prior decision to vote is likely to have reduced the probability of respondents being home or someplace else where they could have heard election news. That would indicate that the relationship exists between hearing news and voting, but the causality is not clear. Most likely, causality goes in different directions for different respondents.

Thus, although intuitively one would expect the micro data would be able to answer the question, we see from an examination of Jackson's careful study that there are major practical problems that, at least to date, have not been resolved. The 1984 Adams study in Oregon that identified nonvoters and interviewed them on the night of the election day to determine reasons for nonvoting seems to be the most powerful micro method yet devised, but unfortunately it studied a landslide election. There are also serious logistical problems in trying to expand this approach to anything approaching a representative national study.

The Macro Level

The alternative method of aggregate analysis has been used widely (Wolfinger and Linquiti, 1981; Tuchman and Coffin, 1971). Most of these papers have looked at a limited number of variables or unsuitable elections, or make unrealistic assumptions. The most realistic paper is by Delli Carpini (1984).

The data are simply aggregate voting data by congressional district. The source of data for Delli Carpini, as well as most others, is statistics summarized by the Congressional Quarterly. To see the effects of the news of exit polls and other election night news in 1980, the percentage of registered voters who voted in 1980 is compared to the same ratio in

Table 2. Surge and Decline in Turnout Rates for the 1980 Presidential and Congressional Races

	Presidential Race	*Congressional Races*
Total districts ($N = 435$)	−2.7% (.3)	−3.8% (.6)
Precall districts ($N = 245$)	−2.1% (.4)	−3.4% (.4)
Postcall districts ($N = 190$)	−3.2% (.4)	−4.4% (.6)
Difference	−1.1%*	−1.0%*

NOTE: Percentages are the mean amount of surge or decline in the respective groupings. Figures in parentheses are standard errors.

* $p \leq .05$.

the 1968, 1972, and 1976 elections for districts where the polls were not and were still open after 8 PM EST when NBC announced the results of their exit survey.

These data are summarized in Table 2. It may be seen that overall there was a decline of 2.7 percent in the voting for president and 3.8 percent in voting for Congress as compared to the average over all four elections. For both races, the difference between the precall and the postcall districts was 1 percent. Not all of this difference may be due to the early call of the election, but the overall impact is small, although not zero.

The macro methods have problems of their own. The first problem is one of aggregation. Della Carpini groups all the states where exit poll results were announced before the polls closed. In 17 of the 28 states where this occurred, the polls closed within an hour of the announcement of results. One would expect that any effects would be reduced in such states. Secondly, Della Carpini did not report on the effect of closeness of the election although he evidently did some analyses controlling for this variable.

Another problem is that other factors may be impacting on turnout. Weather is always a factor and none of the macro analyses control for it. Also the candidates involved influence turnout. Several political commentators have observed that the reduced turnout in California in 1980 was because neither presidential candidate was very popular with the voters. It has been suggested that President Carter's early concession speech may also have been a factor.

Still another problem is that the analyses have usually concerned only the presidential election when the issue of reduced turnout has been raised primarily about congressional races. Della Carpini, however, reports both presidential and congressional races. A final technical issue is that redistricting of congressional districts makes long-term analyses difficult.

To overcome some of these problems an additional analysis was done comparing voter turnout for *close* congressional elections in 1980 with those in 1976, 1978, 1982, and 1984. Close elections were defined as those in 93 districts where the race was decided by less than 15 percentage points in 1980. The states were put into three groups—those where exit poll results were announced after the polls closed, those where they were announced within an hour before poll closing, and those where they were announced more than an hour before closing.

The results are given in Table 3. If one compares voting for Congress in the close districts between 1980 and the other two presidential election years on either side of it, one sees absolutely no difference by the time the exit polls were announced, indicating no net exit poll effects. Indeed, the results suggest a slightly higher turnout overall in these districts in 1980 than for the average of 1976 and 1984.

If, on the other hand, one compares 1980 to the two congressional elections in 1978 and 1982, it may be seen that in districts where the exit polls were announced more than an hour prior to the polls closing the ratio of 1980 to the nonpresidential election years surrounding it was 1.24 while in other districts the ratio is around 1.30, a 6 percentage point difference. Since the standard error of the difference of these ratios is about 7 percentage points, this difference is not statistically different than zero, considering these districts as samples in time of what would occur in similar elections.

It should also be noted that because it has so many districts California dominates these results. If California were excluded there would be

Table 3. Average Congressional Voting per District, 1976–84

	After Polls Close[a]	*One Hour or Less Prior to Polls Close*[b]	*More than One Hour Prior to Polls Close*[c]
1. 1980 vote	192.1	195.1	187.8
2. 1978–82 vote	147.0	150.4	151.5
3. Ratio (1):(2)	1.31	1.30	1.24
4. 1976–84 vote	189.8	192.7	185.5
5. Ratio (1):(4)	1.01	1.01	1.01
(*N*)	(40)	(29)	(24)

[a] Alabama, Connecticut, Delaware, Florida, Georgia, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Mississippi, Missouri, New Hampshire, New Jersey, Ohio, Oklahoma, Pennsylvania, South Carolina, Tennessee, Vermont, Virginia, West Virginia.

[b] Arizona, Arkansas, Colorado, Louisiana, Maine, Michigan, Minnesota, Nebraska, New Mexico, New York, North Carolina, North Dakota, Rhode Island, South Dakota, Texas, Wisconsin, Wyoming.

[c] Alaska, California, Hawaii, Idaho, Iowa, Kansas, Montana, Nevada, Oregon, Utah, Washington.

no differences between the remaining states where the polls closed late and the other states where they closed early.

Thus, depending on the comparison, the differences are zero or 6 percentage points, and if these results are averaged, the average difference is around 3 percentage points.

One issue that is considered by both Jackson and Della Carpini is whether or not there was a differential effect on Democrats or Republicans. Here they reach diametrically opposite conclusions. Jackson concludes that Republicans were most likely to be affected by hearing the early results and that the knowledge of the outcome deters supporters of winners more than supporters of losers. He points out, however, that his results are not statistically significant. Della Carpini concludes, on the other hand, that Republicans were unaffected while the decline in postcall districts was among voters who had early voted Democratic. He points out that his results could simply be caused by the fact that Carter was especially unpopular in the western United States, where polls stayed open later. This explanation has enough reasonability to throw his conclusions into doubt.

One might simply classify congressional districts by whether or not they remained Democratic or Republican in 1980 or whether they switched allegiance. In the close races discussed in Table 3, the percentage of districts that switched from Democrat to Republican in each of the three groups of states arranged by when their polls closed was between 20 and 21 percent. Clearly, there are no switching effects evident for either party.

Summary

As with so many other interesting real-world events, measuring the exact effect of one thing, the exit polls, and controlling for everything else becomes very tricky. Based on a consensus of the data, there is a possibility of a small decrease ranging from 1 to 5 percent in total vote in congressional districts where polls close significantly later than 8 PM EST in those elections where the exit polls suggest a clear winner when previously the race had been considered close. No more precise estimate is possible with the data available.

Whether this is a "significant" effect requiring legislation is not a statistical but a political issue. Most survey researchers will think not, but state and federal legislators ultimately will decide.

References

Adams, William C.
1985 "Early TV calls in 1984: how western voters deplored but ignored them." Paper

presented at 40th Annual Conference, American Association for Public Opinion Research, May 18, 1985, McAfee, New Jersey.

Delli Carpini, Michael X.
1984 "Scooping the voters? The consequences of the networks' early call of the 1980 presidential race." Journal of Politics 46:866–85.

Dubois, Philip L.
1983 "Election night projections and voter turnout in the west." American Politics Quarterly 11:349–63.

Epstein, Laurily K., and Gerald Strom
1981 "Election night projections and West Coast turnout." American Politics Quarterly 9:4 (October): 479–91.

Fuchs, Douglas A.
1966 "Election-day radio-television and western voting." Public Opinion Quarterly 30:226–37.

Jackson, John E.
1983 "Election night reporting and voter turnout." American Journal of Political Science 27:613–35.

Mendelsohn, Harold
1966 "Election-day broadcasts and terminal voting decisions." Public Opinion Quarterly 30:212–25.

Tannenbaum, Percy H., and Kostrich, Leslie J.
1983 Turned-on TV/turned-off voters. Beverly Hills, CA: Sage Publications.

Tuchman, Sam, and Coffin, Thomas E.
1971 "The influence of election night television broadcasts in a close election." Public Opinion Quarterly 35:315–26.

Wald, Kenneth D.
1985 "The closeness–turnout hypothesis; a reconsideration." American Politics Quarterly 13:273–96.

Wolfinger, Raymond, and Peter Linquitti
1981 "Network election day predictions and western voters." Public Opinion 3 (February/March): 56–60.

Television's Impact on High School Achievement

GARY D. GADDY

FROM THE TIME when television was little more than a possibility, observers of the medium have regarded it alternately with hope and with fear. To use E. B. White's description (1938), television seemed likely to be either "an unbearable disturbance of the general peace or a saving radiance in the sky." One particular concern was its potential effect on academic achievement.

When television became reality in the early 1950s, the concern about its effect on children and their school performance mounted. (See, for example, Shayon, 1950.) These fears receded after the early studies failed to produce clear evidence that television was related to achievement.

What little controversy remained at the end of the 1950s was laid to

---

Abstract Using data from the first two waves of the High School and Beyond project, this study examines the reciprocal relationship between media use and academic achievement among American high school students. The hypothesis of the study was that TV influences achievement negatively as it displaces other, more beneficial activities, but that reading influences achievement positively, particularly reading skills. It was also hypothesized that achievement influences media use: specifically, achievement negatively influences television use and positively influences reading. The study used multiple regression in a cross-lagged model. No clear evidence of negative effects of TV on achievement were found. TV, however, was found to be relatively less beneficial than reading for pleasure. And while achievement's influence on the amount of TV viewing was not supported, reading achievement does seem to influence the amount reading.

Gary D. Gaddy is an Assistant Professor at the School of Journalism and Mass Communication and Research Associate in the Mass Communication Research Center at the University of Wisconsin–Madison. This article is based on the author's doctoral dissertation done at the School of Journalism, University of North Carolina at Chapel Hill. An earlier version of this study was presented at the 40th Annual Conference of AAPOR, McAfee, NJ, May 1985. The author would like to acknowledge the substantial contributions of many individuals to this work, especially Jane D. Brown, Rachel Rosenfeld, Eli Rubinstein, Duncan McRae, Jr., Lutz Erbring, Tom Flory, Al Schubert, Gary Pettey, Zena and Frank Biocca, and Doris Graber.

Public Opinion Quarterly Vol. 50:340–359 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-340/$2.50

rest by the parallel findings of two major studies of television and children: Himmelweit et al. (1958) in England, and Schramm, et al. (1961) in America. Both studies found some differences in achievement in favor of children without TV, but they were either statistically insignificant or counterbalanced by other significant differences in favor of children with TV.

A review of the early studies indicates that, although the evidence was weak, the research did suggest that negative effects existed, though on a smaller scale than was first supposed. So, despite—or perhaps because of—the lack of any conclusive evidence, the issue fell dormant. It may be that such benign neglect could be justified in the warm glow of reports of the continually rising national test scores throughout the 1950s and into the 1960s. After all, how could television be so bad when its introduction was followed by almost two decades of academic progress (see Figure 1).[1]

But with the reversal of this trend (see Copperman, 1980), particularly the unprecedented drop in national Scholastic Aptitude Test score averages, the issue was rekindled. A minor irony in this turnabout is that television was raised as a possible contributing factor in this decline by one of the same researchers who had 15 years earlier helped lay it to rest (Schramm, 1976).

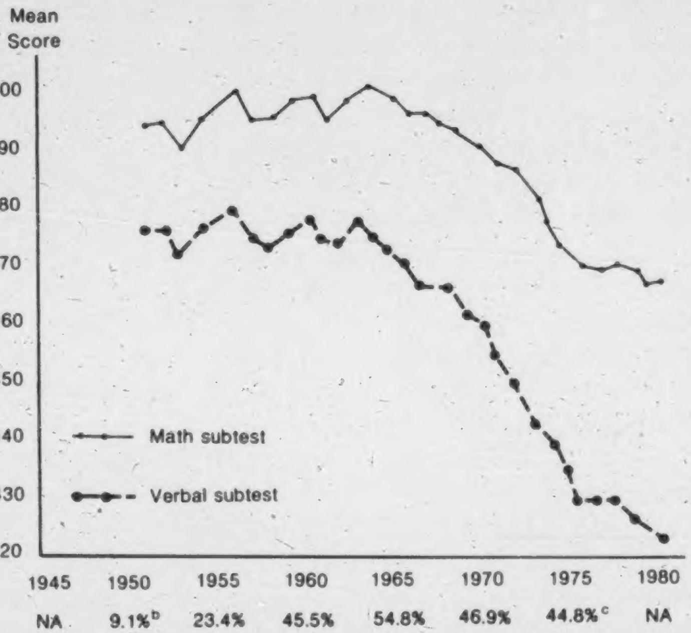
Apparently spurred by this downturn, research on the role of television revived. More recent and generally more sophisticated studies have garnered evidence that television may indeed slow cognitive growth (e.g., Hornik, 1978; Gadberry, 1980). This revival has stimulated several examinations of the state of research itself. While early reviews (e.g., Finn, 1953; Coffin, 1955) were generally inconclusive, more recent summaries (Hornik, 1981; Morgan and Gross, 1982; Comstock, 1982) all support, to varying degrees, the existence of a negative impact of television on achievement. The only clearly dissenting view from this new consensus comes in the form of a news release from an agency of the television industry (Television Information Office, 1982), which appears extremely selective in its reading of the evidence (see Gaddy, 1984: 6–7).

The Process of Effects

While the accumulating evidence suggests a causal effect of television on achievement, the process by which such an effect might take

---

[1] Figure 1 is misleading if it suggests no substantial changes in performance between 1950 and 1965, since during that period the percentage of students taking the test increased from 9 percent to 55 percent. Without overall progress, this increase is likely to have led to a substantial decline in mean scores.



[a] Figure 1 is based on Eckland, 1982, Table 2.1, p. 13. Figures are for all SATs taken in a given year, including those at any grade level, and counting those scores from individuals who repeated the test more than once. Since 1980, SAT scores appear to have stabilized at the bottom of the slide (see Educational Testing Service, 1985, for latest score reports).

[b] Based on 1951–1955 only.

[c] Based on 1975–1976 only.

Figure 1. The Historical Trend in the Scholastic Aptitude Test Scores (SATs) (Annual Means for the Verbal and Math Subtests from 1951–52 to 1979–80)[a]

place is less clear. Hornik's (1981) review of the research outlines six major hypotheses describing how television influences cognitive growth, either positively or negatively. The six hypotheses are: learning of school-equivalent content, interest stimulation, learning of new cognitive skills, learning of instrumental information, development of an intolerance for the pace of schooling, and displacement. While each of these hypotheses has its proponents, they vary in the evidence which may be adduced to support them.

LEARNING OF SCHOOL-EQUIVALENT CONTENT

The hypothesis that television contributes to learning of school-equivalent content is a formal explication of television as a "window on the world." Research has shown that television *can* be an effective teaching tool when it is explicitly designed to do so, as is demonstrated in studies of "Sesame Street" (even the revisionists, Cook et al., 1975, admit that). What has not been demonstrated is whether it *is* an effective teaching tool (Comstock, 1978; Williams, 1981)—as it is presently constituted and used. The evidence that suggests that television has a negative influence on achievement, while not definitive, clearly militates against any hypothesis that suggests it has a positive influence. Only as a source of vocabulary for young children has it consistently been shown to have a positive effect (Schramm et al., 1961; Selnow and Bettinghaus, 1982), and for the latter study this effect appears to be mainly attributable to intentionally educational programs such as "Mister Rogers" and "Sesame Street."

INTEREST STIMULATION

The interest stimulation hypothesis proposes that television stimulates viewers' interest in ideas and topics that they may pursue in forms other than television, particularly book reading. While viewing may stimulate interest in a particular book dramatized on TV (see Splaine, 1978), it does not appear to create an interest in reading, but merely redirects that interest from other reading choices.

The overall impact of television seems to be a reduction in reading, as evidenced by studies of library circulation (Parker, 1963; Tiihonen, 1969). These studies indicate that the introduction of television redirected content (from fiction to nonfiction) and reduced total reading: evidence contrary both to this hypothesis and to the early hopes of the publishing industry (Publisher's Weekly, 1950).

LEARNING OF NEW COGNITIVE SKILLS

The learning of new cognitive skills is a hypothesis that has been advanced most prominently by Salomon (1979), who contends that TV viewing provides practice in the symbol system of television and may allow for the transfer of those skills to the interpretation of other media. While the learning of new cognitive skills has been demonstrated experimentally (e.g., Salomon and Cohen, 1977), extending these findings to more ecologically valid circumstances has succeeded only partially (Huston-Stein and Wright, 1979). Further, it is not clear what impact, if any, these "new cognitive skills" would have on academic achievement, since our educational system is fundamentally print-oriented (Dorr, 1982).

LEARNING OF INSTRUMENTAL INFORMATION

The instrumental information hypothesis says television influences achievement as it molds children's thinking about school, influencing their expectations and aspirations (Berry, 1980). The evidence regarding the association between TV and educational and occupational aspiration, however, is somewhat mixed. Bailyn (1959:29) found high media exposure (TV, movies, and comic books) associated with passive occupational ambitions, but only for boys. Morgan and Gross (1982) found high TV viewing to be associated with lower educational and occupational aspirations for all adolescents one year later, but with higher aspirations for females two years later. Even if aspirations are related to attainment (see Otto and Haller, 1979), the curious and selective pattern of relationships found in the media studies casts doubt on the dependability of these findings.

DEVELOPMENT OF AN INTOLERANCE FOR THE PACE OF SCHOOLING

The hypothesis that television instills an intolerance for the pace of schooling has been promoted by such popular critics of television as Winn (1977). While one study (Anderson et al., 1977) using a brief experimental treatment found no immediate effect of television on tolerance for slow-paced learning, its relevance is limited. As Hornik points out (1981:203), the hypothesis posits accretion, not instantaneous effect. But two other studies (Furu, 1977; Gadberry, 1980), which do address longer-term effects, suggest that television does increase impulsivity, a reasonable operationalization of "an intolerance for the pace of schooling." These studies also suggest that this intolerance is a content effect, not a medium effect (as suggested by Winn, 1977: 33–47), since commercial TV programs were related to increased impulsivity, while noncommercial, "educational" programs were related to increased reflectivity.

DISPLACEMENT

The earliest and most studied explanation for the supposed negative effects of television on achievement is displacement. This hypothesis, paired with the concept of functional equivalence, says that television tends to displace other activities which serve the same needs but are less readily available or require more effort. Further, it argues that television provides less opportunity for "skill practice, contingent reinforcement and active exploration" (Gadberry, 1980:46) than the activities which it displaces, and so inhibits cognitive growth.

The Displacement of Functional Equivalents. Since television consumes 20 hours or more per week for the average teenager (Nielsen, 1981), it must displace something. That television has substantially displaced other "functionally equivalent" entertainment media is incontrovertible. It has substantially displaced cinema, radio, and comic books (e.g., Maccoby, 1951; Furu, 1971).

The most precise estimates of displacement in general come from studies of the use of time. A cross-national comparison of TV owners and nonowners (Robinson, 1972) shows that TV ownership means higher mass media use (including TV, 58 minutes more per day), but lower use of radio (−61%), movies (−52%), books (−41%), and magazines (−28%). Presumably, a substantial part of these changes are due to TV.

The critical question here, however, is not whether television has displaced other activities, but whether or not the displaced activities were more educationally beneficial than the activity (TV) which displaced them. While the displacement of reading would appear to be detrimental to reading achievement, the direct evidence relating to other displaced media is very limited. Only the comparison of television to print seems to provide any direct and clear evidence that TV causes negative displacement (e.g., Wilson, 1974).

Lesser displacement of activities less similar functionally to television have also been noted. For example, studies of TV suggest reductions in time spent in play (Schramm et al., 1961:73), in sleep (Maccoby, 1951; Schramm et al., 1961:73), and also in one activity that would appear to have a direct bearing on academic achievement, homework (Himmelweit et al., 1958:306; Furu, 1971:17). The displacement of homework has been estimated at about 15 minutes a day (Schramm et al., 1961:14; Furu, 1962:21), an amount generally considered to be modest (e.g., Maccoby, 1951).

Displacement as a Mediating Factor. While there is substantial evidence of the displacement effect of television, some suggesting this displacement is negative, little research has directly addressed displacement as a mediating factor in the effect of TV on achievement except for two studies: a longitudinal study (Furu, 1962) and a field experiment (Gadberry, 1980). While Furu's study found no overall relationship between viewing and reading achievement, it did for one of four subgroups, seventh-grade boys—the group which also showed the largest displacement effects for reading and homework. Gadberry's study of first graders showed a significant increase in performance IQ for those children whose TV viewing had been reduced by about half for six weeks. Decreased viewing also led to increased time reading and perhaps coloring: activities positively related to growth in performance IQ. This suggests that for young children television reduces cognitive growth as it displaces more beneficial and ability-enhancing activities.

A General Approach to Displacement. Despite the suggestions of these specific examples of negative displacement, a broader approach to this hypothesis might include a larger definition of displaced activities. Clearly, television displaces a variety of activities—some apparently academically beneficial, others not—most of which individually, or only over the short term, appear to be trivial in consequence. One fallacy of a piecemeal approach to displacement is that the cumulative impact of a "trivial" effect over time may be large. Fifteen minutes of homework may not be much, but 15 minutes per day over 180 school days is 45 hours, and over the course of 12 years of school would total 540 hours—not a trivial amount, especially when this is considered in light of the evidence that "exceptional" cognitive performance can be attained with one hour of practice a day over several years (Walberg et al., 1984).

Further, some of the activity which television does not directly displace it overlays. For example, homework may be reduced directly in amount, or indirectly, when television is watched while homework is done (Maccoby, 1951; Himmelweit et al., 1958:306). Similarly, reading has become more of a "secondary activity" since the introduction of TV (Robinson, 1980:152). If this division of attention reduces the productivity of these activities substantially, then this displacement in quality could be comparable to the more obvious displacement in quantity.

The larger question is not whether the impact of television on one specific behavior has any measurable effect, but whether the total impact across the full spectrum of potentially modified behaviors over time has. Rather than examining only those specific activities presupposed to be displaced, and presupposed to be cognitively beneficial, a more general conception of displacement estimates the overall effect of television mediated through the full range of potentially displaced activities, both beneficial and detrimental. Thus, this framing of displacement acknowledges that television may influence achievement negatively or positively. In other words, the impact of an increase in television viewing will be inversely related to the educational benefit of the activities it displaces.

The overall impact of television on a particular population is hypothesized as the net of the benefit gained while viewing minus the benefit lost from the activities displaced. While television may be helpful to younger children in some respects, it does not appear, in its normal use, to benefit students of high school age or above. If cognitive benefit requires appropriate cognitive challenge, then television will benefit only those who are mentally challenged by it. The relatively low level of effort required by television (particularly as compared to print, Csikzentmihalyi and Kubey, 1981), suggests a declining benefit from television as mental age increases. In contrast, the greater challenge from reading suggests positive effects for reading on achievement, which may not decline with increasing age or mental ability because of the greater "freedom of choice" and "diversity of subject matter" available in print (Himmelweit et al., 1958:36).

Some empirical evidence supports an ability-related decline in benefit from TV. Heavy TV viewing is negatively related to reading achievement for those of medium and high IQ but positively related for those of low IQ (Morgan, 1980), and the negative partial correlation between viewing and reading achievement increases across the elementary grades (Neuman and Prowda, 1982). Similarly, Schramm et al. (1961:274) found declining benefit in vocabulary gain from TV between the first and the tenth grade.

HYPOTHESES

On the basis of the general conception of displacement and the assumption that for the average high school student the sum of the activities displaced by TV is more beneficial to achievement than the TV which displaced it, it may be hypothesized that, *for high school students as a group, television viewing leads to lower achievement.*

Because the effect of displacement is proposed to be directly related to the nature of the activities displaced, it may be hypothesized that *television viewing interacts with the educational environment in its impact on achievement*. Specifically, *the effect of television viewing on achievement will be more negative for those from educationally rich environments than for those from educationally poor environments,* since the probability is greater for the former that the activities displaced will be more beneficial than TV, and for the latter the probability is greater that the activities displaced will be less beneficial than TV.

Reading, in contrast to viewing, is proposed to be beneficial to achievement as it practices a skill fundamental to scholastic achievement—a skill directly measured in many achievement tests. Therefore, it is hypothesized that, *for high school students, reading leads to higher achievement.*

One of the confounding factors in many previous studies of television and achievement has been the failure to consider the impact of achievement on television. Not only may media use influence achievement but achievement may influence media use. Schramm's fraction of selection (1971) proposes that media are selected on the basis of the ratio between the reward offered and the effort required. It follows that if as achievement increases (particularly in reading skills) effort required for reading falls and rewards rise, then *achievement will positively influence the selection of reading*. In addition, because of the relatively minimal effort required by television and the declining benefit it offers as ability increases, it follows that *achievement will negatively influence the selection of television.*

Method

To address the questions posed here, data from the High School and Beyond study (Peng et al., 1981) directed by James Coleman were analyzed. Designed as a national longitudinal study of secondary education, it includes a battery of tests specially created for the assessment of academic achievement. The sample is very large (30,263 sophomores and 28,465 seniors), representative of the United States (drawn using a multistage cluster design stratified at the school level), and currently includes two waves of a panel, collected in 1980 and 1982 with a third wave scheduled for spring 1984 (see National Center for Education Statistics, 1983).

The present study is based upon the analysis of a portion of the 1980 sophomores who were resurveyed in 1982. This sample consisted of students who completed both questionnaires and some of the achievement tests in both waves. Even these respondents number 20,308, so for actual analysis, random samples of this larger sample were taken, leaving subsamples ranging from 2365 to 5074.

MEASURES

Among the vast array of measures on this omnibus survey are a number of tests of academic achievement (comparable to short versions of the Scholastic Aptitude Tests in Figure 1), measures of leisure time use (including television), and an assortment of indicators of personal and social context.

Academic Achievement. The achievement measures include four multiple-choice tests: vocabulary (21 items), reading comprehension (20 items), basic math (28 items), and advanced math (10 items). The same tests were administered in the base year and in the follow-up. In addition, self-reported school grades were analyzed, after being recoded to approximate a college GPA with eight categories ranging from 4.0 ("A's mostly") to 0.5 ("Below D").

Media Use. Television use was measured by answers to the question, "During weekdays about how many hours per day do you watch TV?" Each of the seven categories was assigned its midpoint, with .75 given to "less than 1 hour" and 5.5 to "5 or more." Two other measures of media use, "reading for pleasure" and "reading the front page of the newspaper" were coded according to their approximate weekly frequency in four categories ranging from 0 ("rarely or never") to .75 ("less than once a week") to 4 ("every day or almost every day").

Educational Resources. Indicators of educational resources included a series of potentially educational material possessions:

specifically, having in one's home a daily newspaper, encyclopedia or other reference books, more than 50 books, a room of one's own, and a pocket calculator (all binary coded). In addition, the students' reports of parental educational attainment were included as indicators of personal educational resources, with measures collected in both waves used to increase reliability. The sample was split into high, medium, and low educational resource groups according to the ranking of individuals by an equally weighted sum of material and personal resources, where the weighting of individual items was done using standardized scores.

Other indicators include exposure to education measured by the number of days absent not including illness and by the number of days late to school between the beginning of school and Christmas vacation (both in days), time spent on homework (in hours per week) and how often the student comes to class with homework done (from 0 for "never" to 3 for "always"), being in an academic or college preparatory program and attending a private school (both binary coded). Standard measures of demographics were included as controls, including race, sex, Hispanic origin (binary coded), family income (coded by ordinal rank of the income category), number of parents at home, and number of siblings.

Treatment of Missing Data. Respondents missing data on individual questions were assigned values equal to the mean for the group under analysis (e.g., blacks or females). On family income, respondents missing data on the question with a larger number of categories were assigned according to their answers on a three-category income question if they provided an answer to that question.

STATISTICAL MODEL

The analysis technique used in this study is linear structural modeling of two-wave panel data, a method which is preferable to the older cross-lagged correlation technique, even when correction formulas were used (Kessler and Greenberg, 1981). The primary model was a multiple cross-lagged regression in which the later values of each of the criterion variables were regressed on the earlier values of the criterion as well as the earlier values of the other predictors. This procedure estimates coefficients for each criterion that represent the contribution of each predictor on the change in the criterion, controlling for the contribution of the other predictors.[2]

---

[2] This model simultaneously controls for all the earlier values of all of the variables listed. Exogenous control variables are also included but only as predictors and not as criterion variables.

This analysis makes the usual assumptions necessary for ordinary least squares, that

Results

In the correlations between television and achievement (Table 1), this study clearly replicates the consistent finding of past studies—a uniform and highly significant negative correlation between television and achievement. Further, this relationship holds across the entire series of achievement measures, and for both grade levels.[3] If these correlations were used as a test of our first hypothesis (as has been done), then it would have been overwhelmingly confirmed. Likewise for the second hypothesis, since there is a clear pattern of increasing negative correlations as one moves from the low to the high resource

---

is, that the disturbance terms of the exogenous and endogenous variables are homoscedastic and uncorrelated with each other and any other explanatory variable, and that the equations are linear in their parameters (Marsden, 1981:14–15). Further, it implicitly assumes that the variables are measured without error and that there are no instantaneous effects among the variables. If these assumptions hold, the parameters of this model represent unbiased estimates of causal effect.

Technically, the estimates produced will not be unbiased estimates of the parameters of the cross-lagged effects unless there are no instantaneous effects. Both cross-lagged and instantaneous parameters cannot be simultaneously estimated without very strong assumptions that do not appear to be justifiable in this context. Implicitly assuming no instantaneous effects, as has been done in this analysis, will not lead to estimates biased in direction, unless the instantaneous and lagged effects differ in sign, which seems very unlikely. Further, though the estimates will not be "pure" estimates of parameters of the cross-lagged model specified by Kessler and Greenberg (1981:28–31), they do represent the cumulative effects of the antecedent variables, and in that sense are not biased at all. So, to put it even more technically, the reduced form of the lagged effect is estimated without bias even if the distributed lag effects model is correct.

It is almost certain, however, that none of the assumptions made here are fulfilled entirely. For example, none of the variables contain no error, which means attenuated estimates. Likewise, correlated error between the dependent variables and their lagged values also attenuates estimates of effect (except for the stability coefficients which are inflated). A faulty assumption of linearity also attenuates estimates. A test for nonlinearity in the TV parameter, however, turned out negative.

For control variables, however, measurement error or nonlinearity may attenuate the *reduction* in the spurious relationship. Further, correlated errors (caused by the exclusion of significant "third variables") have no simple and predictable result. For this reason, given the current limited state of knowledge about the processes of achievement and media use, the findings of this study should be viewed as tentative. However, many potential "third variables" have been controlled for in this analysis—including intelligence, implicitly measured by earlier "achievement" scores (see Heyns and Hilton, 1982), and the measures used appear reasonably reliable. So, in the absence of specific, plausible nominations for missing controls, these results should not be summarily dismissed as invalid, particularly since two-stage least-squares results (see Gaddy, 1984) generally confirm those reported here. The pattern of signs was very similar, and although the two-stage least-squares estimates were usually higher in absolute value, they were no more likely to be significant. Along with moderately high cross-temporal correlations (which suggests reasonable stability but some change to be analyzed), this indicates lag misspecification is unlikely to be a great problem.

[3] The zero-order cross-lag correlations do not differ substantially from the simultaneous correlations displayed in Table 2.

Table 1. Zero-order Correlations Between Media Use and Achievement Variables[a]

Television and Achievement—All Students and Subgroups

	All Students		*Blacks*		*Females*	
Achievement Measure	*Soph.*	*Sr.*	*Soph.*	*Sr.*	*Soph.*	*Sr.*
	n = *5074*		n = *2365*		n = *4997*	
Vocabulary	−.158	−.176	−.031	−.023	−.177	−.186
Reading	−.160	−.155	−.042	−.011	−.168	−.160
Math level 1	−.173	−.188	.008	−.011	−.202	−.196
Math level 2	−.157	−.169	−.021	−.030	−.159	−.176
Grades	−.128	−.127	−.032	.021	−.146	−.136

Television and Achievement—Educational Resource Groups

	High Resource		*Medium Resource*		*Low Resource*	
Achievement Measure	*Soph.*	*Sr.*	*Soph.*	*Sr.*	*Soph.*	*Sr.*
	n = *2974*		n = *3960*		n = *3112*	
Vocabulary	−.241	−.220	−.076	−.121	−.037	−.094
Reading	−.226	−.174	−.064	−.089	−.064	−.086
Math level 1	−.211	−.204	−.105	−.122	−.084	−.109
Math level 2	−.171	−.175	−.097	−.110	−.073	−.113
Grades	−.189	−.148	−.080	−.089	−.051	−.074

Media Use and Achievement—All Students

	Viewing Television		*Reading for Pleasure*		*Reading Newspaper*	
Achievement Measure	*Soph.*	*Sr.*	*Soph.*	*Sr.*	*Soph.*	*Sr.*
	n = *5074*		n = *5074*		n = *5074*	
Vocabulary	−.158	−.176	.228	.266	.164	.147
Reading	−.160	−.155	.245	.269	.136	.135
Math level 1	−.173	−.188	.155	.160	.128	.112
Math level 2	−.157	−.169	.136	.126	.083	.094
Grades	−.128	−.127	.173	.133	.133	.115

[a] Smallest significant correlation (one-tailed alpha = .05, corrected for design effect) for all students is .032, for blacks .047, for females .032, for high resource group .042, medium resource group .036, and for low resource group .041. Calculated using Fisher's *r* to *Z* transformation.

group. In addition, the negative correlations for females would replicate the finding for all students, and the insignificant correlations for blacks would appear to confirm that those relatively deprived educationally are less "harmed" by television viewing.

The potentially misleading nature of purely correlational data is indicated by the results of a more appropriate analysis. As Table 2 indicates, the cross-lagged, standardized partial regression coefficients (controlling for a range of exogenous demographic, social, and educational characteristics as well as earlier levels of the various measures of achievement, media use, and educational activities) provide little support for the hypothesis that television is generally detrimental in the sense that its use directly reduces achievement growth. Although there

Table 2. Cross-lagged Coefficients—Television as Independent Variable (Including All Control Variables)[a]

Dependent Variable	*All students Beta n = 5074*	*Blacks Beta n = 2365*	*Females Beta n = 4997*
Vocabulary achievement	−.004	.015	−.004
Reading achievement	−.001	.006	.006
Mathematics level 1	−.012	.012	−.014
Mathematics level 2	−.009	−.017	−.017
Grades	−.008	−.008	−.003

[a] In addition to the variables listed above, each cross-lagged regression equation includes: sex, race, hispanicity, private school attendance, academic program, father's and mother's education, number of siblings and parents at home, family income, educational possessions (5 measures), homework (two measures), absenteeism and tardiness, frequency of reading for pleasure and the front page of the newspaper. Race and sex excluded in the respective subgroup regressions.

All underlined coefficients are significant at .05 (one-tailed corrected for design effect).

is a relatively consistent pattern of negative coefficients for both the total sample and for females, the estimated "effects" are small and in no case statistically significant. The estimates for blacks are mixed between small negative and positive coefficients.

Similarly, Table 3 provides no support for the hypothesis that the negative effect of television increases with the level of educational resources. If anything, the pattern suggests the opposite, but note that none of the differences between coefficients are statistically significant.

Table 3. Cross-lagged Coefficients—Television as Independent Variable. Sample Divided by Level of Educational Resources (Including All Control Variables)[a]

Dependent Variable	*High Resource Beta n = 2974*	*Medium Resource Beta n = 3960*	*Low Resource Beta n = 3112*
Vocabulary achievement	−.001	.023	−.012
Reading achievement	.011	.015	−.013
Mathematics level 1	.000	.000	−.004
Mathematics level 2	−.014	−.017	−.022
Grades	−.002	−.009	−.013

[a] In addition to the variables listed above, each cross-lagged regression equation includes: sex, race, hispanicity, private school attendance, academic program, father's and mother's education, number of siblings and parents at home, family income, educational possessions (5 measures), homework (two measures), absenteeism and tardiness, frequency of reading for pleasure and the front page of the newspaper.

All underlined coefficients are significant at .05 (one-tailed, corrected for design effect).

Also, the only significant coefficient is one for the medium resource group and it is positive. Again, although the majority of coefficients are negative, no support for the first hypothesis can be adduced since none of the negative coefficients is statistically significant. A replication of this test, comparing blacks (considering them as a relatively deprived subgroup) to the total population, again provides little support for the displacement hypothesis (see Table 2).

A comparison of TV viewing, reading for pleasure, and reading the front page of the newspaper, however, does indicate that reading for pleasure is positively related to reading and vocabulary achievement (see Table 4). Further, the differences between reading for pleasure and TV viewing, and between reading for pleasure and reading the front page of the newspaper, are statistically significant. This suggests that viewing television and reading the front page of the paper are relatively unbeneficial when compared to reading for pleasure.

The hypothesized negative effect of achievement on television viewing is only marginally supported. Even though four out of five coefficients are negative only one is statistically significant (see the bottom half of Table 4). The hypothesized positive impact of achievement on print use receives fairly clear support, at least for the types of achievement most closely related to reading. Vocabulary achievement appears to influence both reading for pleasure and reading the front page of the newspaper, and reading achievement to significantly influence reading for pleasure. The significant negative coefficient for the influence of math level one on reading for pleasure seems somewhat incongruous, but consider that this estimate is after controlling for the other achievement measures which are all highly, positively related.

Discussion

There are two major sources of explanation for the diminished negative relationship between television and academic achievement: reverse causation and "third variables." The bottom half of Table 4 (and similar results for the subgroups) suggests that part of this diminution is due to the effect of achievement on the selection of television. Although these "reverse" effects are neither large nor consistently significant, they are generally negative and larger than the effects from TV to achievement.

The greater part of the observed correlation appears to be due to the effect of "third variables," which is to say, variables apparently external to the relationship between television and achievement. That is, differences in such personal and social characteristics as race, sex, parental education, educational resources, and intelligence (implicitly

Table 4. Cross-lagged Coefficients (Total Sample Including All Control Variables)[a]

	Media Use as Independent Variables		
Dependent Variable	Viewing Television Beta	Reading for Pleasure Beta	Reading Newspaper Beta
Vocabulary achievement	−.004	.047	.011
Reading achievement	−.001	.082	.008
Mathematics level 1	−.012	.003	−.006
Mathematics level 2	−.009	.004	−.010
Grades	−.008	−.011	.014

	Media Use as Dependent Variables		
Independent Variable	Viewing Television Beta	Reading for Pleasure Beta	Reading Newspaper Beta
Vocabulary achievement	.005	.124	.067
Reading achievement	−.026	.055	−.008
Mathematics level 1	−.058	−.049	−.038
Mathematics level 2	−.023	.008	−.032
Grades	−.016	−.021	.049

NOTE: For all equations $n = 5074$.

[a] In addition to the variables listed above, each cross-lagged regression equation includes: sex, race, hispanicity, private school attendance, academic program, father's and mother's education, number of siblings and parents at home, family income, educational possessions (5 measures), homework (two measures), absenteeism and tardiness.

All underlined coefficients are significant at .05 (one-tailed, corrected for design effect).

measured by earlier achievement test scores) may influence both television viewing and achievement, creating an appearance of a substantial causal relationship that does not hold under closer analysis.

Several cautions are in order regarding the conclusions that may be drawn from these estimates. While the fairly consistent failure to obtain statistically significant effects does mean that the null hypothesis of no effects cannot be rejected, it does not mean that it demonstrates that television has no effect on achievement. Clearly, it casts doubt on previous studies which found evidence of negative effects on American high school students but used fewer controls or less adequate methods. Just as clearly, it does not refute the evidence of other studies which also found evidence of negative effects on other populations and used alternative methods of equal or greater validity, for example, Gadberry's (1980) field experiment with first graders or Hornik's (1978) quasi-experimental investigation in El Salvador.

Even in the presence of statistical significance, these findings might be considered evidence that television has no socially significant impact because of the minute size of the estimated effects. Given the number of conservative biases built into the statistical procedures used in the study, especially the assumption of perfect measurement and a procedure that in effect uses change scores (see Kessler and Greenberg, 1981:141), a more reasonable approach is to compare effects within the study before passing such a judgment.

One logical basis for comparison is the set of estimates for other media activities, such as reading for pleasure or reading the front page of the newspaper, and "obvious" sources of academic effects such as high school program, private school attendance, and time spent on homework. The estimated effect of TV (measured by betas averaged across the four achievement tests for the total sample) does turn out to be the only negative estimate (–.007), but is not radically different in absolute value from the other estimates: reading the newspaper (.001), private school attendance (.019), homework time (.021), reading for pleasure (.034), and being in an academic program (.042). A comparison across an average of the effects on the two verbal measures, vocabulary and reading comprehension, offers a similar picture: TV (–.003), newspaper (.010), private school (.031), homework (.014), reading for pleasure (.065), and academic program (.041).

Clearly, television does not fare well in these comparisons—and the size of all of the estimates suggests caution in defining a meaningful effect in absolute terms. But even if the major suggestion of this study does hold, that is, that television viewing has no measurable effect on achievement, this is considerably less than a glowing endorsement of television as an instrument of education. In comparison to reading for pleasure, for example, television appears clearly less effective.

To whatever degree this study supports the hypothesis that television has no negative effect on achievement, it even more clearly supports the hypothesis that it offers no benefit. In other words, it suggests that the thousand hours a year high school students invest in television are, as Comstock (1982:1154) put it, "scholastically unproductive."

IMPLICATIONS FOR FUTURE RESEARCH

The major shortcoming of this study is the measure of television viewing. As the authors of a recent study (Roberts et al., 1984) point out, operationalizing television viewing only in terms of gross amount ignores many aspects of the phenomenon, including involvement, social context, and the content of what is viewed. While a global measure of viewing such as was used in this study does address the significant question of the overall impact of television, that is, viewing cumulated across levels of involvement, varieties of social contexts, and types of content, it does not address many more specific ones. When possible, future studies should.

Even if television has, in the aggregate, no measurable effect, it would be both socially and theoretically meaningful to know whether this is because regardless of how, where, or what is viewed there are no effects, or because counterbalancing effects tend to neutralize each other. For example, is the relative benefit of those who watch to learn offset by the relative loss of those who watch to escape? Or is the positive effect of viewing news, public affairs, and science programming balanced out by the negative effects of entertainment programming?

Even though at the social level, the net effect is the same regardless of the answers to these questions, a fuller understanding of the processes by which media influence cognitive development clearly depends upon their careful investigation. Such an understanding may serve us well as we enter an age with a continually changing array of new hopes and fears concerning the latest addition to the panoply of new communications technologies.

References

Anderson, Daniel R., Stephen R. Levin, and Elizabeth P. Lorch
1977 "The effects of TV program pacing on the behavior of preschool children." A-V Communication Review 25:159–66.
Bailyn; Lotte
1959 "Mass media and children: a study of exposure habits and cognitive effects." Psychological Monographs: General and Applied 73 Whole No. 471:1–48.
Berry, Gordon L.
1980 "Children, television and social class roles: the medium as an unplanned educational curriculum." Pp. 71–81 in Edward L. Palmer and Aimee Dorr (eds.), Children and the Faces of Television: Teaching, Violence, Selling. New York: Academic Press.
Coffin, Thomas E.
1955 "Television's impact on society." The American Psychologist 10 (October):630–41.
Comstock, George
1978 Trends from the Study of Incidental Learning from Television Viewing. Syracuse, NY: ERIC Clearinghouse on Information Resources.
1982 "Mass media." Pp. 1149–57 in Harold E. Mitzel (ed.), Encyclopedia of Educational Research (5th ed.). New York: The Free Press.
Cook, Thomas D., Hilary Appleton, Ross F. Conner, Ann Shaffer, Gary Tamkin, and Stephen J. Weber
1975 "Sesame Street" Revisited. New York: Russell Sage Foundation.

Copperman, Paul
1980 "The decline of literacy." Journal of Communication 30 (Winter):113–22.
Csikszentmihalyi, Mihaly, and Robert Kubey
1981 "Television and the rest of life: a systematic comparison of subjective experience." Public Opinion Quarterly 45:317–328.
Dorr, Aimee
1982 "Reading of the age of television." Paper presented at the Annual Meeting of the Claremont Reading Conference, January 14–15, Claremont, CA (ERIC Reproduction Service No. ED 216 310).
Eckland, Bruce K.
1982 "College entrance examination trends." Pp. 9–27 in Gilbert R. Austin and Herbert Garber (eds.), The Rise and Fall of National Test Scores. New York: Academic Press.
Educational Testing Service
1985 National College-bound Seniors, 1985. New York: College Entrance Examination Board.
Finn, James D.
1953 "Television and education: a review of research." Audio-Visual Communication Review 1 (Spring):106–26.
Furu, Takeo
1962 Television and Children's Life: A Before-After Study. Tokyo: Japan Broadcasting Corporation.
1971 Function of Television for Children and Adolescents. Tokyo: Sophia University.
1977 "Cognitive style and television viewing patterns of children." Research report. Department of Audio-Video Education, International Christian University (June).
Gadberry, Sharon
1980 "Effects of restricting first graders' TV-viewing on leisure time use, IQ change, and cognitive style." Journal of Applied Developmental Psychology 1:45–57.
Gaddy, Gary D.
1984 "Television and scholastic achievement: a study of American high school students." Ph.D. dissertation University of North Carolina at Chapel Hill.
Heyns, Barbara, and Thomas L. Hilton
1982 "The cognitive tests for High School and Beyond: an assessment." Sociology of Education 55:89–102.
Himmelweit, Hilde T., A. N. Oppenheim, and Pamela Vince
1958 Television and the Child. London: Oxford University Press.
Hornik, Robert C.
1978 "Television access and the slowing of cognitive growth." American Educational Research Journal 15:1–15.
1981 "Out-of-school television and schooling: hypotheses and methods." Review of Educational Research 52 (Summer):193–214.
Huston-Stein, Aletha, and John C. Wright
1979 "Children and television: effects of the medium, its content and its form." Journal of Research and Development in Education 13:20–31.
Kessler, Ronald C., and David F. Greenberg
1981 Linear Panel Analysis: Models of Quantitative Change. New York: Academic Press.
Maccoby, Eleanor E.
1951 "Television: its impact on school children." Public Opinion Quarterly 15:421–44.
Marsden, Peter V.
1981 "Concerning single equation models." Pp. 13–21 in Peter V. Marsden (ed.), Linear Models in Social Research. Beverly Hills: Sage Publications.
Morgan, Michael
1980 "Television viewing and reading: does more equal better?" Journal of Communication 30 (Winter):159–65.

Morgan, Michael, and Larry Gross
1982 "Television and educational achievement and aspiration." Pp. 78–90 in David Pearl, Lorraine Bouthilet, and Joyce Lazar (eds.), Television and Behavior: Ten Years of Scientific Progress and Implications for the Eighties (Vol. 2): Technical Reviews. Washington, DC: U.S. Government Printing Office.
National Center for Education Statistics
1983 High School and Beyond: 1980 Sophomore Cohort First Follow-up (1982)—Data File User's Manual. Washington, DC: National Center for Education Statistics (April).
Neuman, Susan B., and Peter Prowda
1982 "Television viewing and reading achievement." Journal of Reading 25 (April):666–70.
Nielsen, A. C.
1981 Child and Teenage Television Viewing. Nielsen Television Index.
Otto, Luther B., and Archibald O. Haller
1979 "Evidence for a social psychological view of the status attainment process: four studies compared." Social Forces 57 (March):887–914.
Parker, Edwin B.
1963 "The effects of television on library circulation." Public Opinion Quarterly 27:578–89.
Peng, Samuel S., William B. Fetters, and Andrew J. Kolstad
1981 High School and Beyond: A National Longitudinal Study for the 1980's: A Capsule Description of High School Students. Washington, DC: U.S. Department of Education (April).
Publisher's Weekly
1950 "Books and TV." Publisher's Weekly 157 (June 17):2638–44.
Roberts, Donald F., Christine M. Bachen, Melinda C. Hornby, and Pedro Hernandez-Ramos
1984 "Predictors of reading achievement at different age levels." Communication Research 11 (January):9–49.
Robinson, John P.
1972 "Television's impact on everyday life: some cross-national evidence." Pp. 410–31 in Eli A. Rubinstein, George A. Comstock, and John P. Murray (eds.), Television and Social Behavior (Vol. 4): Television in Day-to-Day Life: Patterns of Use. Washington, DC: U.S. Government Printing Office.
1980 "The changing reading habits of the American public." Journal of Communication 3 (Winter):141–52.
Salomon, Gavriel
1979 Interaction of Media, Cognition, and Learning: An Exporation of How Symbolic Forms Cultivate Mental Skills and Affect Knowledge Acquisition. San Francisco: Jossey-Bass.
Salomon, Gavriel, and Akiba A. Cohen
1977 "Television formats, mastery of mental skills and the acquisition of knowledge." Journal of Educational Psychology 69:612–19.
Schramm, Wilbur
1971 "Nature of communication between humans." Pp. 3–53 in Wilbur Schramm and Donald F. Roberts (eds.), The Process and Effects of Mass Communication (rev. ed.). Urbana: University of Illinois Press.
1976 Television and the Test Scores. Princeton: College Board and Educational Testing Service.
Schramm, Wilbur, Jack Lyle, and Edwin B. Parker
1961 Television in the Lives of Our Children. Stanford: Stanford University Press.
Selnow, Gary W., and Erwin P. Bettinghaus
1982 "Television exposure and language development." Journal of Broadcasting 26 (Winter):469–79.
Shayon, Robert Lewis
1950 "The Pied Piper of video." The Saturday Review of Literature, November 25, pp. 49–51.

Splaine, John E.
1978 "Television and its influence on reading." Educational Technology 18 (June):15–19.
Television Information Office
1982 "Television and education: assessing academic achievement." TIO News Release (December 1).
Tiihonen, Pertti
1969 "Influence of television on the use of libraries." Report 7, Finnish Broadcasting Company.
Walberg, Herbert J., Bernadette F. Strykowski, Evangelina Rovai, and Samuel S. Hung
1984 "Exceptional performance." Review of Educational Research 54 (Spring):87–112.
White, E. B.
1938 "One man's meat." Harpers Magazine, October, pp. 553–56.
Williams, Tannis Macbeth
1981 "How and what do children learn from television?" Human Communication Research 7 (Winter):180–92.
Wilson, C. Edward
1974 "The effect of medium on loss of information." Journalism Quarterly 51 (Spring):111–15.
Winn, Marie
1977 The Plug-in Drug. New York: Viking Press.

Education and the Quality of Life in Korea and the United States: A Cross-Cultural Perspective

DOH C. SHINN

> Men possess a moral nature; but if they are well fed, warmly clad, and comfortably lodged, without being taught at the same time, they become almost like the beasts.
>
> —MENCIUS

IN RECENT YEARS, an increasing number of social scientists in the United States and elsewhere have made serious efforts to determine whether or not more education contributes to the quality of people's lives (Andrews, 1985; Andrews and Withey, 1976; Atkinson, 1978; Campbell et al., 1976; Campbell, 1981; Gallup, 1976; Glenn and Weaver, 1981; Headey, 1981; Inglehart, 1977; Inglehart and Rabier, 1985; Rabier, 1974; Veroff et al., 1981). Unfortunately, these research efforts have

---

Abstract The search reported in this article analyzed two sets of national sample survey data recently collected from the United States and the Republic of Korea, and sought to compare the influence of educational attainment on the perception of life quality in these two countries. Results of this comparison indicate that education contributes a lot more to the psychological quality of life in Korea than in America. Two explanations were offered for the different roles which educational attainment plays in the two countries.

Doh C. Shinn is a Professor of Political Studies and Public Affairs at Sangamon State University in Illinois. His research on the quality of life was supported by a postdoctoral research grant from the Social Science Research Council and a travel grant from the Association for Asian Studies. The data used in this article were made available by the Inter-University Consortium for Political and Social Research, University of Michigan, and the Institute of Social Sciences, Seoul National University. For their valuable advice at all stages of his work on life quality, the author thanks Professors Ahn Chung-Si, Chey Myung, Chung se-koo, Khil Seung-Hem, Kim Kyong-Dong, and Lee Hongkoo at Seoul National University, Dr. Conrad Rutkowski at Fordham University, and Professors Craig Brown and Jack Van Der Slik at Sangamon State University. He also gratefully acknowledges the helpful comments and suggestions of the three anonymous reviewers on an earlier version of this article and the assistance of Jackie Wright in preparing tables, figures, and text.

Public Opinion Quarterly Vol. 50:360–370 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-360/$2.50

been confined to advanced industrialized societies in Western Europe and North America; very little systematic research has been conducted to compare the effects of education on psychological well-being across culturally and economically divergent types of societies.

The present study seeks to fill this void in the psychological study of life quality by analyzing two sets of national sample survey data recently collected from the Republic of Korea and the United States. Unquestionably, America represents one of the most affluent and educationally advanced countries in the West. Korea, on the other hand, is one of the most rapidly developing, nonwestern countries of the Third World. More important, Korean and American societies have long been shaped by different cultures, Confucianism and Christianity. These differences in their economic wealth, educational attainment, and cultural heritage provide an interesting opportunity to explore the dynamics of education in the psychological world of human well-being.

Does more education help people in America and Korea to live happier, more pleasant, more satisfying, and more successful lives? Or does educational attainment increase aspirations and expectations so much that it makes it more difficult to experience life as "better"? In which of these two countries does education have stronger and more pervasive effects on psychological well-being, and why?

Conceptualization and Measurement

The concept of quality of life has been defined and measured in many different ways. McCall (1975), for example, viewed it as a life of happiness (see also Shinn and Johnson, 1976). Other scholars have equated life quality with a sense of satisfaction (Atkinson, 1978; Campbell et al., 1976). Still others have related it to the extent to which life goals are accomplished (Hankiss, 1980; Milbrath, 1982). Undoubtedly, such unidimensional approaches to the concept of life quality portray a partial picture of the quality of life perceived by people themselves.

This study measures quality of life in terms of four widely known elements of the good life: happiness, enjoyment, satisfaction, and accomplishment (von Wright, 1972). Happiness and enjoyment represent the hedonic dimension of the good of man, while satisfaction and accomplishment constitute its welfare dimension. By considering at least these two dimensions of life quality, we are attempting to make a more comprehensive and balanced analysis of education's effects on it.

In measuring educational attainment, the present study focuses on learning experience in public and private schools. Unquestionably, there are important differences in what Korean and American schools teach and how they teach their students. Yet, these differences are minor compared to those between other types of educational institutions in the two societies. The Korean school system since World War II has been modeled after the American system. Both systems require six years of schooling in elementary education, another six years in high school education, four years in college education, and at least two years in postgraduate training. In view of the parallels, we measure educational attainment in terms of six mutually exclusive categories ranging from no formal school to college graduation.

Two sets of personal and environmental characteristics are controlled for statistically in order to estimate the independent effects of education on the subjective quality of life accurately. The first set includes gender, age, marital status, and family income. The second set features the types of housing and community in which Korean and American respondents live.

The variable of gender is dichotomized into male and female. Marital status was also dichotomized into married and the unmarried by collapsing those widowed and divorced into the unmarried category. Family income is divided into quintiles. Age was aggregated into five categories: 18–29, 30–39, 49–49, 50–59, and 60 and older. The variable of housing is measured on the basis of whether a unit was owned by its residents or not. Community is measured on an ordinal scale based on the size of its population. The scale incudes three categories: large metropolitan areas with populations of over 500,000, medium and small-size cities with populations of 25,000 to 500,000, rural towns and isolated villages.

Data Collection and Analysis

Basic data for the present study came from two nationwide sample surveys on the quality of life, which were carried out separately in two different periods of time. The first survey was conducted by the Institute of Social Sciences at Seoul National University during the summer of 1980 (Shinn et al., 1982). The second was conducted by the Survey Research Center at the University of Michigan during the summer of 1978, two years before the Korean survey. Besides this difference in the timing of field work, the number of households chosen for personal interviews differed considerably between the surveys. The sample size was 1215 for the Korean survey and 3642 for the American survey. This difference in sample size made it impossible to carry out many interesting comparisons of the various subgroups of the two populations. The two samples, however, are sufficiently representative to explore systematically cross-cultural differences in the influences of education on psychological well-being.

Included in both surveys were more than 60 comparable items on perceptions of well-being and ill-being, education and many other personal and environmental characteristics. Of these items, we selected four questions designed to tap the general feelings of well-being and seven items dealing with education and other relevant characteristics of respondents.

In measuring a general sense of satisfaction, both surveys used the same 7-point rating scale, which ranged from complete dissatisfaction to complete satisfaction. As a measure of life enjoyment, they also asked respondents the same question, "How much do you enjoy your life these days—a great deal, a fair amount, not very much, or not at all?" In measuring the extent to which happiness was experienced and life goals were accomplished, the Korean survey offered more response categories than the American survey had done earlier.

In an attempt to minimize the possible effects that these differences in question wording may have had on the reported levels of well-being, we first collapsed the answers of all respondents into the positive and negative categories. For each of four general items on life quality, we then computed the proportions of Koreans and Americans who gave positive answers. Finally, for each sample we computed the average number of positive answers and the proportion who answered all four questions positively. The mean scores of positive replies were used to compare the overall levels of life quality experienced by the adult populations of the two countries.

Four individual and two composite measures of life quality were cross-classified with levels of educational attainment. This bivariate analysis highlights the magnitude of differences between the more and less educated of the two populations in their experiences of well-being. In addition, multiple classification analysis (MCA) was performed on the same measures in order to estimate the independent effects of education on well-being and its relative importance as a source of subjective quality of life. Unlike other techniques of multivariate analysis, MCA is capable of estimating the value of the dependent variable for each category of independent variable after the adjustment of all other explanatory variables in the equation (Andrews et al., 1973).

Findings

Table 1 presents the percentages of Korean and American respondents who replied that they were happy, that they enjoy their lives, that they were satisfied with their lives, and that they had accomplished what they wanted in their lives, categorized by level of educational

Table 1. Percent of American and Korean Respondents Expressing Feelings of Well-Being, by Educational Attainment

Measures of Well-Being	Educational Attainment Levels						
	No School	*Elementary School*	*Middle School*	*High School*	*Some College*	*College Grad*	*Entire Sample*
America							
Happiness	—[a]	85%	87%	91%	95%	92%	91%
Enjoyment	—	73	76	85	88	89	84
Satisfaction	—	75	77	80	84	85	81
Accomplishment	—	87	88	85	87	89	86
All four	—	32	38	43	43	45	41
(*N*)		(178)	(533)	(364)	(875)	(371)	(3456)
Korea							
Happiness	42%	54%	55%	58%	63%	75%	57%
Enjoyment	38	48	54	51	50	65	51
Satisfaction	14	32	44	50	51	65	43
Accomplishment	22	32	28	30	41	51	32
All four	5	11	12	15	21	33	15
(*N*)	(88)	(292)	(246)	(364)	(92)	(104)	(1092)

[a] Not enough cases.

attainment. On every measure of well-being, Korean adults are shown in the table to rate their lives far less positively than Americans.

Comparing the ratings of individual measures, we see that the size of Korean adult population expressing positive feelings is from 31 percent to 54 percent smaller than that for the American adult population. Comparing the ratings of composite measure reported in the same table, we find that Korean people, as compared to their American counterparts, are two and a half times less likely to experience all the feelings of happiness, enjoyment, satisfaction, and accomplishment (15 versus 41 percent). Such wide differences in the general level of well-being expressed by the two populations are not surprising in view of the economic, political, and other environmental conditions in which they live.

It is surprising that general levels of well-being vary across educational categories a great deal more in Korea than in America. In America, college graduates are found to express positive feelings about their lives from 2 to 16 percent more than those whose education ended in elementary school. In Korea, however, the differences between the same educational groups run from 12 to 33 percent. Judged by the percentage expressing all four feelings of well-being, college graduates in America appear to be better off less than one-half the time when compared to their counterparts with just an elementary level of education (45 percent versus 32 percent). In Korea, on the other hand, college degree holders are found to be three times better off psychologically than those with only an elementary education (33 percent versus

11 percent). It appears that a life of well-being has more to do with educational attainment in Korea than in America.

To estimate educational differences in the experiences of well-being more accurately, statistical adjustments were made for the effects of six control variables on those experiences. Table 2 reports levels of well-being for six educational categories after the original scores were adjusted to compensate for differences in family income, gender, age, marital status, housing, and community. As expected, educational differences were reduced considerably in both samples. The average rates of such reduction, however, are found to be far greater for the American sample than for the Korean sample. This indicates that income and five other personal and environmental characteristics exert a relatively more powerful influence on the quality of life in America than in Korea.

From Table 2, we can see that American respondents' adjusted scores of well-being vary little from one educational category to another on the measures of happiness, satisfaction, and accomplishment. Only in the assessment of life enjoyment do we observe a statistically significant difference (12 percentage points between the two extreme categories of educational attainment). Between Korean respondents in the same categories, however, the differences run from 22 to 44 percent, and these differences are all found to be statistically significant and substantively meaningful. One obvious conclusion from this

Table 2. Educational Differences in the Adjusted Scores of Well-Being for American and Korean Respondents

	Educational Attainment Levels						
Measures of Well-Being	*No School (A)*	*Elementary School (B)*	*Middle School (C)*	*High School (D)*	*Some College (E)*	*College Grad (F)*	*Difference (F − D)*
America							
Happiness	—[a]	89%	89%	91%	94%	91%	0%
Enjoyment	—	76	77	85	87	88	3
Satisfaction	—	79	80	80	83	84	4
Accomplishment	—	89	88	85	85	89	4
All four	—	33	38	42	42	44	2
(*N*)		(178)	(533)	(364)	(875)	(371)	(3456)
Korea							
Happiness	48%	54%	55%	58%	63%	70%	22%
Enjoyment	40	48	54	51	49	62	11
Satisfaction	17	34	45	48	48	61	13
Accomplishment	25	32	28	30	39	47	17
All four	3	11	12	16	22	33	17
(*N*)	(88)	(292)	(246)	(364)	(92)	(104)	(1092)

[a] Not enough cases.

finding is that, with more education, Koreans experience a much better quality of life than their American counterparts do.

The differences in the role which education plays in Korean and American societies become clearer when we compare the adjusted ratings of two composite measures. The number of positive feelings expressed by Korean respondents varies a great deal from 1.2 for those with no formal schooling to 2.6 for those with college degrees even when their family income and other important characteristics were statistically controlled. Among American respondents, in sharp contrast, the same measure is found to vary only by 0.2, from 3.0 to 3.2, a variation which is hardly meaningful at all. According to the adjusted ratings of the percentage scale, Koreans with college degrees are 11 times more likely to experience a life of happiness, enjoyment, satisfaction, and accomplishment than those with no formal education (33 versus 3 percent). College graduates in America, on the other hand, are only slightly more likely to experience it than the least educated (44 versus 33 percent). On the basis of these findings, one can reasonably argue that education contributes far more to the quality of life in Korea than in America.

The greatest difference attributable to education in the two countries appears to be in the way college graduation affects the perceptions of life quality. To estimate the net effect of a college degree on life quality, we subtracted the adjusted scores of the high school educated from those of college degree holders in each country. Among those in the Korean sample, over two times as many college graduates as the high school educated reported that they were living a life of happiness, enjoyment, satisfaction, and accomplishment (33 versus 16 percent). The margin of this difference for the American sample was no more than 2 percent (42 to 44 percent). On the whole, the possession of a college degree appears to be far more essential to the experience of life quality in Korea than in America.

All the evidence presented above indicates strongly that education is a more powerful and positive influence on the psychological quality of life in Korea than in America. The next logical question is this: How much more powerful an influence and how many more positive effects does educational attainment have on the quality of life when contrasting Korea with America? To answer this question, we performed multiple regression analyses on the overall measure, which summed up the number of positive answers to four questions on well-being. Table 3 reports simple correlation coefficients, regression coefficients, and standardized regression coefficients (beta) for education and six other explanatory variables.

Table 3. Relative Importance of Education as a Source of the Perceived Quality of Life in America and Korea

Explanatory Variables	*Simple Correlation Coefficients*		*Regression Coefficients*			
			Unstandardized		*Standardized*	
	America	*Korea*	*America*	*Korea*	*America*	*Korea*
Education	.10*	.22*	.07*	.16*	.07*	.17*
Income	.26*	.28*	.07*	.26*	.10*	.25*
Age	−.02*	−.06*	.01	−.04	.06	−.08
Gender	.05	.01	.05	−.09	.03	−.03
Marriage	.14*	.06	.16*	.43*	.09*	.12
Housing	.13*	.12*	.14*	.22*	.07*	.07*
Community	.09*	−.04	.07*	−.22*	.06*	−.15*

* $p < .05$.

A comparison of regression coefficients across the columns of the table reveals that one unit increase in educational attainment alone would improve the quality of life among Koreans by a unit of 0.17. Among American people, on the other hand, the same unit of change in educational attainment would raise their life quality by a unit of 0.07. This means that education shapes the psychological quality of life in Korea two and a half times more powerfully and positively than in America.

Comparing betas across the rows of the table, we are able to identify the relative importance of education as a source of life quality in each country. In both countries, we see that income ranks first among all seven variables in terms of the value of the beta statistic. For Americans as well as Koreans, income appears to be the most important determinant of their life quality. This finding is hardly surprising in view of the widely known fact that monetary resources nowadays not only fulfill survival needs but also purchase personal pleasure, power, prestige, and status.

The most surprising result of this comparison is that education ranks differently as a source of life quality in the two countries. In Korea, education ranks second, right after family income. In America, however, education ranks fourth, after marital status and place of residence. Indeed, education as a source of life quality occupies a far more important place in Korea than in America. More investments in education may not contribute much to public well-being in America, but such investments appear to be essential to the building of a nation of well-being in Korea.

Discussion

Having found the significant differences in the way education influences the psychological quality of life in Korea and America, we now need to explore the reasons for these differences. Obviously, one of these reasons has to do with differences in the extent to which education is emphasized as a life goal in the two countries. People in Korea, like the Chinese and Japanese, whose culture has been dominated by Confucian ethics, still believe strongly that a good quality of life cannot be attained without higher education (Dawson, 1942: chaps. 1–2). For this reason, children in these countries even attempt to terminate their precious lives when they fail to be admitted to the school of their choice. For the same reason, newly arrived Asian-Americans are represented far beyond their population share at virtually every top-ranking university in this country. Despite their minority status, they are also known for achieving higher socioeconomic levels than the white majority (Hirschman and Wong, 1984).

A great deal of difference also exists between Korea and America in the way education is associated with other tangible resources known to be necessary for a good quality of life. In the two countries, family income is distributed far more closely in line with levels of education in Korea. According to the percentage statistics reported in Figure 1, the probability that college graduates become a member of the highest income group in their society is more than two and a half times greater

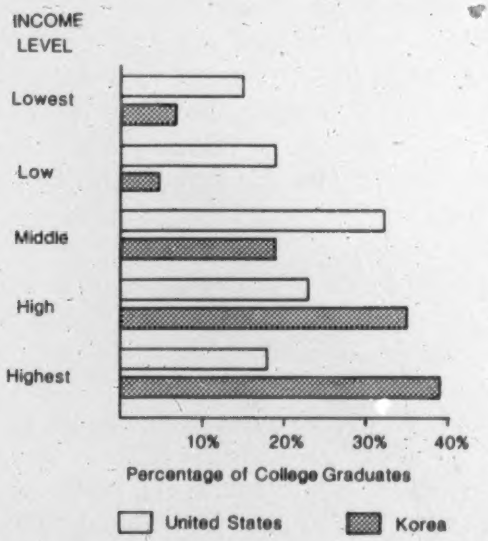


Figure 1. Percentage of U.S. and Korean Respondents with a College Degree in Five Levels of Income

in Korea than in America (40 versus 17 percent). Therefore, college graduation itself will be psychologically more significant to Koreans than Americans. This may be another reason why education serves as a more powerful and positive force influencing the quality of life in Korea than in America.

Conclusions

Education is thought to be an important means by which we add to our quality of life. Each year, millions of people, both in developing and developed countries, search for more and better educational opportunities in public or private schools. In prior research, however, few efforts have been made to compare the effects of educational achievement on the perception of quality of life across countries whose economies and cultures are at variance. The research reported here analyzed two sets of national sample survey data collected from the Republic of Korea and the United States, and sought to assess and explain the dynamic roles which educational attainment plays in psychological well-being. All the findings documented above indicate strongly that education contributes a lot more to the public's well-being in Korea, where the Confucian concept of self-development through education is still widely shared:

> Those who are born with the possession of knowledge are the highest class of men. Those who learn and so acquire knowledge are the next. Those who are dull and stupid, and yet compass the learning, are another class next to these. As to those who are dull and stupid and yet do not learn, they are the lowest of the people (Confucious, *Analects,* Book XVI, C. IX).

References

Andrews, F. M. (ed.)
1985 Studies in Subjective Well-Being. Ann Arbor: University of Michigan.
Andrews, F. M., and S. B. Withey
1976 Social Indicators of Well-Being. New York: Academic Press.
Andrews, F. M., et al.
1973 Multiple Classification Analysis. Ann Arbor: University of Michigan.
Atkinson, T.
1978 Public Perceptions of Quality of Life in Canada. Toronto: Institute of Behavioral Research, York University.
Campbell, A.
1981 The Sense of Well-Being in America. New York: McGraw-Hill.
Campbell, A., P. E. Converse, and W. L. Rodgers
1976 The Quality of American Life. New York: Russell Sage Foundation.
Confucious
1971 Confucian Analects (translated by J. Legge). New York: Dover Publications.

Dawson, M.M.
1942 The Basic Teachings of Confucious. New York: New Home Library.
Gallup, G.
1976 "Human needs and satisfaction: a global survey." Public Opinion Quarterly 40:22–39.
Glenn, N., and C. Weaver
1981 "Education's effects on psychological well-being." Public Opinion Quarterly 45:22–39.
Hankiss, E.
1980 The Roles of Goals in People's Lives. Paris: UNESCO.
Headey, B.
1981 "The quality of life in Australia." Social Indicators Research 9:151–181.
Hirschman, C., and M. G. Wong
1984 "Socio-economic gains of Asian Americans, blacks, and Hispanics: 1960–76." American Journal of Sociology, 90:584–607.
Inglehart, R.
1977 The Silent Revolution. Princeton: Princeton University Press.
Inglehart, R., and J. Rabier
1985 "If you are unhappy, this must be Belgium: well-being around the world." Public Opinion 8:10–15.
McCall, S.
1975 "Quality of life." Social Indicators Research 2:229–48.
Mencius
1970 The Works of Mencius (translated by J. Legge). New York: Dover Publications.
Milbrath, L.
1982 "A conceptualization and research strategy for the study of ecological aspects of quality of life." Social Indicators Research, 10:135–58.
Rabier, J. R.
1974 Satisfaction et Insatisfaction Quant aux Conditions de Vie dans les pays Membres de la Communaute Europenne. Brussels: Commission of the European Committee.
Shinn, D. C., C. Ahn, K. Kim, and H. Lee
1982 "Perceptions of quality of life in an industrializing country: the case of Korea." Social Indicators Research 10:475–92.
Shinn, D. C., and D. Johnson
1976 "Avowed happiness as an overall assessment of quality of life." Social Indicators Research 5:475–92.
Veroff, J., E. Douvan, and R. Kulka
1981 The Inner American: A Self-Portrait from 1957 to 1976. New York: Basic.
Von Wright, G. H.
1972 The Variety of Goodness. London: Routledge and Kegan Paul.

Response Effects in the Electronic Survey

SARA KIESLER AND LEE S. SPROULL

IN ALL SCIENTIFIC ENDEAVOR the available tools affect the questions one can ask and the data one can collect. In this report, we examine a new tool in survey research, the electronic or computer-mediated survey. In the last two decades, electronic computers have come to figure in many phases of survey research—instrument design, sampling, monitoring of work in the field, coding and editing, data entry, data cleaning, scale and index construction, data base organization, data base retrieval, statistical analysis, documentation, and report writing (Kar-

---

Abstract This report examines the electronic survey as a research tool. In an electronic survey, respondents use a text processing program to self-administer a computer-based questionnaire. As more people have access to computers, electronic surveys may become widespread. The electronic survey can reduce processing costs because it automates the transformation of raw data into computer-readable form. It can combine advantages of interviews (e.g., prompts, complex branching) with those of paper mail surveys (e.g., standardization, anonymity). An important issue is how the electronic survey affects the responses of people who use it. We conducted an experimental sample survey on health attitudes, behaviors, and personal traits using two forms of administration: electronic and paper mail. Closed-end responses in the electronic survey were less socially desirable and tended to be more extreme than were responses in the paper survey. Open-ended responses that could be edited by respondents were relatively long and disclosing. These findings are consistent with other research on computer-mediated communication, raising general issues about using computers to collect self-report data.

Sara Kiesler is a Professor of Social Sciences and Social Psychology, and Lee Sproull is an Associate Professor in the Department of Social and Decision Sciences, Carnegie-Mellon University. This research was supported by grants from the National Science Foundation, and from IBM and the System Development Foundation to the Committee on Social Science Research in Computing at Carnegie-Mellon University. Elise Yoder, Andrew McFarlane, Robert Hines, Mark Wisniewski, Alexander Czajkowski, Michael Tarr, and Jane Siegel helped design the procedures and collect and analyze the data. Some of the survey items were taken from a longitudinal study of health attitudes and behavior by Richard Lau. The electronic survey program used in this research was written by Mike Blackwell. Kathleen Carley, Joel Greenhouse, James March, Allen Newell, Charles Turner, and two anonymous reviewers provided valuable comments on previous drafts of this report.

Public Opinion Quarterly Vol. 50:402–413 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-402/$2.50

weit and Meyers, 1983:379–414). Still, the full application of computers to survey research is limited by the fact that raw survey data are collected off-line. The technology in common use that most closely approximates the electronic survey is computer-assisted telephone interviewing (CATI). In CATI, the interviewer types oral responses directly into the computer, which improves the efficiency of data collection over oral-response-to-paper-to-computer methods. However, in CATI, a person (the interviewer-recorder) stands between the respondent's behavior and the data that are analyzed. By contrast, in an electronic survey respondents at computer terminals or microcomputers enter their own data directly into the computer.

The sampling frame of the electronic survey is restricted to members of organizations and populations who have access to computers and to people who feel comfortable using them. Until such time as computers and networks spread throughout society, the electronic survey will probably be infeasible for general surveys. More limited applications, however, are increasingly appropriate. Greist and Klein and their colleagues have used free-standing microcomputers to collect diagnostic information directly from psychiatric patients (e.g., Greist et al., 1973). Researchers could use free-standing or portable computers in shopping centers and convention centers to carry out market studies, in military bases and firms to carry out personnel research, and in schools and government agencies to carry out program evaluations. The data collected on small computers can be transferred by telephone to any central data processing source.

When people are linked by a computer network, the electronic survey can be administered to any person having access to the network or to a connected network. Kraut (1984) used a computer network in an international corporation to conduct research on work behaviors and attitudes. Sproull (1985) used computer mail in a Fortune 500 firm to study organizational communication. Computer networks add value to the electronic survey because these systems locate respondents automatically, deliver survey instruments to remote locations (wherever respondents have access to the network), and permit respondents to answer questions at their own convenience. Computer networks also make feasible automatic event-driven surveys, for example, a consumer survey keyed automatically to particular purchases by shoppers who are using electronic catalogs.

If only because it seems to reduce research costs, the electronic survey may become widespread. Once respondents have access to a computer or to a network, relatively lower marginal costs of collecting and communicating data electronically can be substituted for the substantial costs of interviewing, telephoning, and sending questionnaires through the mail. In addition, 20–40 percent of the total computing costs of conducting a survey may derive from transforming data collected off-line into a form which can be processed by the computer (Ferrara and Nolan, 1974:27). The electronic survey eliminates the need for a person (such as a transcriber) or a technology (such as an optical character reader) to stand between respondents and the computer that stores and analyzes their data. Some open-ended material would still have to be hand-coded, but even this task could be aided by computerized text search and text categorizing programs. Another potential advantage of the electronic survey is that the instrument can be standardized easily, as is true of a paper questionnaire. But as with the telephone or face-to-face interview, the instrument is highly flexible. It may have any style, formal or informal. It can include explanatory material, prompts, error corrections, menus, branches, and skips. Because of this flexibility, the electronic survey is potentially adaptable to many kinds of research including organizational case studies, political polls, attitude surveys, experiments, and evaluations.

Assessing Survey Utility

Apart from any economic advantages, the utility of the electronic survey will depend on its comparability to other methods of survey administration. This comparability is not obvious because the electronic survey both shares characteristics of various other methods and has some unique features of its own. One way to approach this issue is through a comparison of nonsampling errors, that is, of variations from theoretically true survey responses. After Bradburn (1983:289) and others, we will refer to these errors as "response effects."[1]

Every survey administration introduces or exacerbates response effects (DeLamater, 1982:37). These include respondents' systematically refusing to answer certain questions, or giving incomplete answers or not following instructions, underreporting socially undesirable or threatening information, overreporting socially desirable information, choosing conventional or "moderate" response categories, and "yea-saying"—agreeing with whatever the researcher asserts. Face-to-face and telephone interviews increase respondents' desire to please over self-administered paper questionnaires. Hence interviews increase the quantity of responses and detail in reports of behavior; they also increase overreporting of socially desirable attributes or attitudes (for a secondary analysis and review, see Sudman and Bradburn, 1974). Self-administered questionnaires are relatively anonymous and tend to reduce respondents' concern over presenting themselves in a good light. Hence self-administered questionnaires reduce total reporting and accuracy but increase reporting of negative information and attitudes (e.g., Kahn, 1952; Hochstim, 1967).

---

[1] We use the term *response effects* because we did not compare subjective responses with objective data and therefore cannot establish the degree of nonsampling error in the strict sense of that term. We recognize that this expediency avoids deeper measurement issues surrounding the nature of true subjective information.

Because the essence of both an electronic survey and the self-administered paper questionnaire is answering questions presented via printed text, we expect responses to an electronic survey to be much like responses to a paper survey. However, like the telephone and face-to-face interview, the electronic survey is interactive. Interactive features such as prompts might increase respondents' perceived control and attention, with the result of increasing their involvement beyond that evoked by a paper questionnaire.

The electronic survey is also unique because of the degree to which it lacks social context information. All survey administration methods, because they standardize interaction, reduce social context information beyond that which we gather in most interactions in everyday life. Mail surveys reduce social context information more than telephone surveys do, telephone surveys more than face-to-face interviews. However, electronic surveys, because they lack even the social artifacts embodied in letterheads, postal rate markings, and print formats reduce social context information most of all.[2] Research suggests that reduced social context information in the electronic survey will make the research setting impersonal and anonymous, and that respondents will become self-centered, and relatively unconcerned with social norms and with the impression they give others (e.g., Short et al., 1976; Kiesler et al., 1984).

Economic considerations will probably make electronic surveys popular for some kinds of survey research. On the basis of our ideas about technologically induced shifts of attention and reduced social context information, we predict specific differences in response effects for the electronic survey in comparison with the paper mail survey. The remainder of this report presents the results of our first experiment.

Method

During the fall of 1983, we conducted an experiment to compare responses in the electronic survey with those in an equivalent paper survey. On the basis of the reasoning described above, we formulated two sets of working hypotheses about relative response rates and response effects.

---

[2] The social quality of electronic communication is not simply a matter of narrow bandwidth. Developers of computer communication programs and text editors deliberately design these programs to maximize flexibility and freedom from the social constraints embodied in other methods of communication.

Absorption/attention effects:

Respondents in an electronic survey will complete more items and will complete them with fewer mistakes than will respondents in a paper survey.
Respondents in an electronic survey will give longer answers to open-ended questions than will respondents in a paper survey.
Respondents in an electronic survey will talk more about themselves than will respondents in a paper survey.

Lack of social context information effects:

Respondents in an electronic survey will less frequently choose "agree" responses and middle responses in Likert scales than will respondents in a paper survey.
Respondents in an electronic survey will admit to more socially undesirable behavior, and they will report more personal traits and feelings, than will respondents in a paper survey.

SETTING

This research was conducted at Carnegie-Mellon University, a private university enrolling 6100 students. Students and employees have wide access to computers; more than 85 percent of the students have a computing account. Every liberal arts student is required to take an introductory programming course (usually PASCAL) and many other courses make use of the computer. The most frequently used computer programs, in and out of classes, are the communication and text-editing programs run on a network of time-sharing computers (Blackwell, 1984). Hence, the experiment was conducted in a relatively computer-intensive environment, where many people use computers to communicate text.

SAMPLE SELECTION AND STUDY PROCEDURES

On the basis of a pilot study we conducted in spring 1983, we decided to sample from the population of recently active computer mail users. On each of three days and on each of eight computers, we ran a system program that lists all currently logged-in users and the name of the program they are running. (At any given time, about one in six users are running a mail program.) We also invoked a system command to list every person who has a visible mail file, along with the date and time it was last used. From the first 300 computer mail users identified by these methods, we randomly selected 76 students and 75 faculty/staff employees. This resulted in a sample roughly proportionate to the distribution of employees and students across colleges and administrative units, with a small oversampling of employees in Computer Science, Robotics, and the Computation Center (18.7 percent versus 14.8 percent in the university as a whole).

We sent each person in the sample a letter inviting him or her to participate in a research project on electronic surveys. The invitation explained that the respondent would receive either a hard-copy or an on-line version of a questionnaire drawn from current social science projects. It also promised that respondents' data would be anonymous and would be used for statistical purposes only. Follow-up telephone calls were made to the group to ask verbal permission to proceed with the survey. Overall, 93 percent received the invitation (141/151) and 71 percent of the group that received the invitation (100/141) agreed to participate and actually returned the questionnaire.

SURVEY INSTRUMENT

The survey instrument we used was a simple 18-item self-administered questionnaire on health and personal characteristics. Fifteen of the items were closed-ended: five questions were used to group respondents by gender, organizational position, and computer use; five 7-point Likert scales on attitudes about health (e.g., "There is nothing more important than good health") were used to measure yea-saying and extreme responses; five true-false questions from the Marlowe-Crowne (1964) Need for Approval Scale were used to measure social desirability responses. Three of the items were open-ended: One asked the respondent to describe his or her most recent illness; one asked, "What are the habits and reactions of yours which bother you at present?" One asked, "What characteristics of yourself give you cause for pride and satisfaction?"

The paper survey was prepared with a high-quality laser printer. The electronic survey was a computer program which the respondent invoked at his or her terminal to display each question and prompt for answers. The program allowed respondents to retype any answer after they had finished the questionnaire, but it did not permit respondents to edit their answers as they completed the questionnaire.

When the respondent had completed the electronic questionnaire he or she returned it to a master storage directory on one computer. There, numerical identification codes were added automatically. Then the coded data (without names) were forwarded to a different computer directory where data analyses were performed. This procedure protected the anonymity of respondents' data while also permitting us to contact respondents in the future by matching codes to mailing label information in the storage directory.

PRELIMINARY ANALYSES

Correlations among the dependent variables indicate that our primary measures were independent (r's = .01–.03). There were some group effects—females wrote longer open-ended responses, and heavier computer users gave less socially desirable answers—but group variables did not interact with survey method effects.

Results

PARTICIPATION IN THE SURVEY

More respondents returned the paper survey than returned the electronic survey (75 percent vs. 67 percent). Still, the electronic response rates compare favorably with rates of 48–61 percent in conventional mail surveys (Heberlein and Baumgartner, 1978). Furthermore, the electronic survey was returned more quickly than was the paper survey (9.6 days vs. 10.8 days).

Overall, the survey respondents consisted of 84 percent male students and 54 percent male employees, which corresponds to 70 percent male students in the university as a whole and 57 percent male (noncustodial) employees in the university as a whole. Respondents reported spending 3.7 hours computing on a "typical" weekday (which is not very different from what was reported by the pilot study sample). Male and female respondents reported an equal number of hours for computer use. Gender and time spent using the computer did not affect response rates or response times.

ITEM COMPLETION

Respondents who answered the electronic survey made fewer item completion mistakes (e.g., writing an integer in response to a true-false question) on the closed-ended items than did respondents who answered the paper survey. Of the total of 53 items where a mistake had been made in answering a question (5.3 percent of the total possible correct), none was made by elecronic survey respondents. Electronic survey respondents also left fewer items blank and refused to answer fewer questions than did paper survey respondents. Of the total of 20 missing or refused items, 8 came from electronic survey respondents. Overall, 22 percent of the paper survey respondents failed to complete or spoiled one or more items, whereas only 10 percent failed to complete or spoiled one or more items in the electronic survey.

OPEN-ENDED ANSWERS

Total number of words, of personal pronouns, and of trait descriptions did not differ significantly between the electronic survey and the paper survey. Several respondents in the electronic survey condition sent us computer mail complaining about the absence of text editing capabilities in the electronic survey. Later in this report we describe some data relevant to editing of open-ended items.

RESPONSE SETS

We used the health attitude Likert scales to check two kinds of response set: the tendency to agree (yea-saying), and the tendency to avoid extreme answers, that is, to prefer the midpoints of scales. Since the format for each item was a 7-point scale where 1 = "agree" and 7 = "disagree," smaller mean scores indicate more yea-saying. To estimate moderation in responding, we subtracted 4 (the midpoint of the scale) from the respondent's answer; a smaller absolute difference indicates a less extreme answer. On the basis of an analysis of variance to measure the effects of survey method, position (staff or student), and item, we found no significant differences in yea-saying or extremity of response attributable to survey method.

SOCIAL DESIRABILITY EFFECTS

To measure the social desirability of responses we used items from the Need for Approval Scale (Marlowe and Crowne, 1964). Published norms suggest that most people answer in a socially desirable direction; in all cases this is a "true" response. Our data indicate that respondents who answered electronically gave less socially desirable responses than did respondents who answered on paper. Analysis of variance on the factors in Table 1 (survey method, position, and items) showed a significant main effect for survey method, $F(1,87) = 3.8$, $p < .05$, and no interactions.

EXTENSION

Four months after we completed the experimental survey, we mailed letters to the 100 respondents (using a mailing list generated from the storage directory) asking them to participate in an extension to the electronic survey study. Sixty-three people volunteered. One purpose of our follow-up was to explore the reliability of our findings by giving respondents the instrument they had not received previously. We did so with 33 of the volunteers; 20 of the (previous) electronic survey

Table 1. Percentage of Socially Desirable Responses

	Method of Administration			
	Students		Employees	
Item	*Paper* (n = *26*)	*Electronic* (n = *25*)	*Paper* (n = *25*)	*Electronic* (n = *24*)
I am always careful about my manner of dress.	50%	50%	52%	20%
I always try to practice what I preach.	75	67	79	72
When I don't know something I don't at all mind admitting it.	80	75	61	60
I would never think of letting someone else be punished for my wrongdoings.	90	92	87	57
I never resent being asked to return a favor.	74	63	74	59
Mean %	74%	69%	71%	54%

respondents completed the paper version, and 13 of the (previous) paper survey respondents completed the electronic version.

Although the number of respondents is small, we found differences similar to those we had observed earlier or even stronger in the predicted direction: (1) New electronic survey respondents were more cooperative than were new paper survey respondents (faster response time, more items completed, fewer mistakes in responding). (2) New electronic survey respondents were less likely to be "yea-sayers" ($F[1,24] = 4.1, p < .05$) and more likely to use extreme scale points than were paper survey respondents ($F[4,120] = 2.64, p < .05$). (3) New electronic survey respondents gave less socially desirable answers than did new paper survey respondents ($F[1,21] = 3.41, p < .08$).

Another purpose of our follow-up was to collect data on response effects in an open-ended instrument that could be edited freely by respondents. We sent 30 of the study extension volunteers the three open-ended items that had appeared in our experimental survey as electronic questions that could be freely edited. An example of one edited response follows:

> I get excited too easily and sometimes get impatient with people and react too spontaneously when it would sometimes be better to "go with the flow." I also feel that I ought to give up alcohol, caffeine and smoking my pipe, he said as he took a puff.
>
> There are also days when I feel that I push too hard, both myself and the people

who work for me. And I sometimes worry that I demand too much and don't encourage people enough.

Respondent-edited answers in the brief questionnaire were twice as long as were responses to the same items in the paper survey and more than three times as long as were responses in the standard electronic survey, $F(2,49) = 7.55$, $p < .01$. (The statistics reflect comparisons among all study extension volunteers.) Respondent-edited responses also were more self-centered, as measured by the number of first-person personal pronouns in relation to third-person personal pronouns, $F(2,49) = 3.52$, $p < .05$ (e.g., Davis and Brock, 1975). Finally, respondent-edited responses included more self-description, as measured by counting personal attributes and traits in the answers ($F[2,49] = 5.0$, $p < .01$). The standard electronic survey, however, elicited the most trait descriptions per words written.

Discussion

The population of interest for an electronic survey will be a community or organization with access to and familiarity with computers or computer networks. These groups will tend to be relatively well-educated, urban, white collar, and technologically sophisticated. Assuming the electronic survey is feasible with respect to this population, it seems to offer some advantages over a paper survey. The results of our experiment suggest that the electronic survey, at least one administered within an organized setting, can elicit good response rates with faster turnaround time and fewer item incompletions than a regular mail survey. Our survey was simple, but it need not have been. The electronic survey can be designed to handle any of many formats prohibited in self-administered paper surveys: open-ended questions requiring long answers, corrections, questions which screen and then branch to other questions, and questions which must be answered in certain sequences (Dillman, 1983:359, 369–370).

Our results show considerable similarity of response between the paper and electronic survey but not so much that the two may be considered interchangeable without further research. Even in our small sample we found more socially undesirable responses in the electronic survey than in the paper survey. In other research we have found that people tend to be both self-absorbed and uninhibited when they communicate using a computer (Kiesler et al., 1984). Sproull (1985) found that employees answering a questionnaire using electronic mail gave more extreme answers than employees who answered on a paper questionnaire. Within the context of response effects this work suggests a caution and counterweight to economic factors that encourage the use of electronic surveys. (See, for example, Green, 1983, on personality testing via computer, Slack, 1971, on medical history-taking via computer, and Knight et al., 1980, on performance appraisal via computer.) At a minimum, electronic survey results should be calibrated against those obtained through more conventional means. Beyond that, electronic survey results may give us a glimpse into how people behave in a new communication environment.

References

Blackwell, Mike
1984 "Using the computer to automatically collect data about user behavior." Committee on Social Science Research in Computing Working Paper, Carnegie-Mellon University.

Bradburn, Norman M.
1983 "Response effects." Pp. 289–328 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.

Crowne, Douglas, and David Marlowe
1964 The Approval Motive. New York: Wiley.

Davis, Deborah and Timothy C. Brock
1975 "Use of first person pronouns as a function of increased objective self-awareness and prior feedback." Journal of Experimental Social Psychology 11:381–88.

DeLamater, John, and Kathleen McKinney
1982 "Response-effects of question content." Pp. 13–48 in Will Dijkstra and Johannes van der Zouwen (eds.), Response Behavior in the Survey-Interview. London: Academic Press.

Dillman, Don A.
1983 "Mail and other self-administered questionnaires." Pp. 359–77 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.

Ferrara, Raymond, and Richard L. Nolan
1974 "New look at computer data entry." Pp. 25–46 in William C. House (ed.), Data Base Management. New York: Petrocelli Books.

Green, Bert F.
1983 "Computer testing." Letter to the editor, Science (December 16) 222:4629.

Greist, John H., Marjorie H. Klein, and Harold P. Erdman
1976 "Routine on-line psychiatric diagnosis by computer." American Journal of Psychiatry 12:1405–08.

Heberlein, Thomas A., and Robert Baumgartner
1978 "Response rates to mailed questionnaires." American Sociological Review 43:447–62.

Hochstim, Joseph R.
1967 "A critical comparison of three strategies of collecting data from households." Journal of the American Statistical Association 62:976–89.

Kahn, Robert L.
1952 "A comparison of two methods of collecting data for social research: the fixed-alternative questionnaire and the open-ended interview." Unpublished doctoral dissertation, University of Michigan.

Karweit, Nancy, and Edmund D. Meyers, Jr.
1983 "Computers in survey research." Pp. 379–414 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research. New York: Academic Press.

Kiesler, Sara, Jane Siegel, and Timothy McGuire
1984 "Social psychological aspects of computer-mediated communication." American Psychologist, 39:1123–34.

Knight, James L., Leslie A. Geddes, and Gavriel Salvendy
1980 "Equipment note: continuous, unobtrusive performance and physiological monitoring of industrial workers." Ergonomics 23:501–06.

Kraut, Robert
1984 Telecommuting: Cautious Pessimism. Murray Hill, NJ: Bell Communications Research.

Short, John, Ederyn Williams, and Bruce Christie
1976 The Social Psychology of Telecommunications. London: John Wiley & Sons.

Slack, Warner
1971 "Computer-based interviewing system dealing with nonverbal behavior as well as keyboard responses." Science (January 8) 171:84–87.

Sproull, Lee
1985 forthcoming. "Using electronic mail for data collection in organizational research." Academy of Management Journal.

Sudman, Seymour, and Norman M. Bradburn
1974 Response Effects in Surveys. Chicago: Aldine.

The Polls:
Medical Care in the United States

ROBERT Y. SHAPIRO AND JOHN T. YOUNG

MEDICAL care and government policies toward it have been changing slowly but significantly in the United States. Rapidly rising medical costs—especially cost increases larger than high general inflation during difficult economic times—became the catalyst for change. The Reagan administration's proposals, policies, and budget cuts have altered the course of health policy. Problems of equity and access to health care continue for the poor and uninsured (estimates of the latter ranging from 8 to 15 percent), and the recurring debate about forms of national health insurance may begin anew (cf. Erskine, 1975; Kingdon, 1984). Other related issues which also affect the costs of medical care are starting to become more prominent: the needs of an aging and older population, long-term care, catastrophic care, prevention, the development and use of new technology, and malpractice and the accountability of the medical profession. The public is only beginning to get acquainted with the new ways (and alphabet soup) in which patient care is provided, insured, and paid for—larger insurance premiums, deductibles, and out-of-pocket expenses; and, for example, the uses of preferred provider organizations (PPOs), health maintenance organizations (HMOs), and the new prospective payment cost controls for treatments: the diagnostic-related groups (DRGs) that are currently applied to Medicare and inevitably more broadly.

Where, then, does the public stand? Public concern and support for government assistance in medical care is virtually on par with Social Security as an *entitlement* (cf. Shapiro and Smith, 1985). This support appears to have been strikingly stable for decades, and on the basis of limited data, it seems to be strikingly similar to public attitudes in other affluent Western countries that have more extensive national medicare programs (cf. Coughlin, 1980; Pescosolido et al., 1985). Support for more spending and government action in this policy area has been consistently high in general, especially because of the elderly and the poor, who are perceived to be truly needy. The public appears

---

This is the second in a series of articles on public opinion and social welfare policy.

Robert Y. Shapiro is Assistant Professor of Political Science at Columbia University and a NORC Research Associate. John T. Young is a doctoral student in Political Science at Columbia. The authors thank Mary L. DeFlorio, M.D., Tom W. Smith, and Sasha Heid for their comments and assistance. Support has been provided by the Ford Foundation Project on Social Welfare Policy and the American Future and by the Columbia University Council for Research in the Social Sciences; the responsibility for analysis and interpretation is the authors'.

Public Opinion Quarterly Vol. 50:418–428 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-418/$2.50

to be motivated by compassion and altruism toward *all* citizens, although some broadly defined private interests cannot be ruled out (cf. Steiber and Ferber, 1981; Shapiro and Smith, 1985).

Overall, the survey data suggest that the changes in health care thus far have not had substantial effect on the public's general satisfaction with health care and medical insurance (but cf. John, 1985). However, in addition to challenging high costs, citizens appear to be more critical, litigious, and scrutinizing than in the past of the behavior of doctors and hospitals. Systematic evidence concerning the consequences of various reforms and policy changes has not yet been examined, but the public, given the medical services that it has come to expect, will tend to oppose drastic changes in access or quality such as rationing (e.g., Aaron and Schwartz, 1984), and abuses or widespread (and highly publicized) errors in judgment that occur in the name of "cost effectiveness" (cf. Bernard, 1985, on uncompensated hospital care and patient "dumping," and recent anti-dumping legislation).

On the other hand, there appears to be an underlying ambivalence that makes the public very receptive to more incremental changes of the sort that occurred in the past with Medicare, Medicaid, federal encouragement or support for HMOs and certain types of special treatments (e.g., renal dialysis), and, currently, prospective payment controls and changes in employer-provided health plans (see Taylor, 1985: 19–20). Such policies along with other credible and persuasive efforts at opinion leadership could move the country closer, with uncertain consequences, to some form of national health insurance (cf. Morone and Dunham, 1985; Sidel and Sidel, 1984; Page et al., 1987).

National health insurance has generally had a good deal of support in principle, but this has not necessarily been support for *governmentally* provided insurance or health services, but rather the desire that the federal government guide private and public efforts to provide and pay for medical care (see Payne, 1946). Americans are ever willing to look to market competition for assistance—particularly a market tending toward a surplus of medical providers—but with typically desirable government regulation in the wings (e.g., Shapiro and Gillroy, 1984).

The survey questions reported here were taken from the files of the Policy Change and Public Opinion Change Project at NORC (originally supported by the National Science Foundation, grant #SES-7912969 A01). These data were compiled from published and unpublished sources, and with the assistance of the Roper Center, University of Connecticut; the Louis Harris Data Center, University of North Carolina at Chapel Hill; the Inter-university Consortium for Political and Social Research, University of Michigan; and the NORC Library. Further information about these data can be obtained from the appropriate archive or survey organization. The national sample sizes are at least 1000 unless noted otherwise.

CRI — Cambridge Reports, Inc., including Cambridge Reports (1985)
HARRIS — Louis Harris and Associates
HARRIS/KFF — HARRIS for the Kaiser Family Foundation
KPA/AMA — Kane, Parsons, and Associates, Inc., for the American Medical Association
MOR — Market Opinion Research
NBC — National Broadcasting Company
NYT/CBS — *The New York Times*/Columbia Broadcasting System

NORC — NORC, University of Chicago
NORC-GSS — NORC's General Social Surveys
ROPER — Roper Organization
SRC/CPS — Survey Research Center/Center for Political Studies, University of Michigan, American National Election Studies
TRENDEX — Trendex, Inc./General Electric Quarterly Index of Company Relations. NB: Samples drawn from telephone directories, not RDD
YSW — Yankelovich, Skelly, and White, Inc.
YSW/HIAA — YSW for the Health Insurance Association of America

References

Aaron, Henry J., and William B. Schwartz
1984 The Painful Prescription: Rationing Hospital Care. Washington, D.C.: The Brookings Institution.

Bernard, Bruce P.
1985 "Private hospitals dumping of patients." The New York Times, October 28, p. A19.

Cambridge Reports, Inc.
1985 America Looks Ahead. Cambridge, MA: Cambridge Reports, Inc.

Coughlin, Richard
1980 Ideology, Public Opinion, & Welfare Policy. Berkeley, CA: Institute of International Studies.

Erskine, Hazel
1975 "The polls: health insurance. Public Opinion Quarterly 39: 128–41.

John, Kenneth E.
1985 "Americans' satisfaction with their health care is on the decline." The Washington Post National Weekly Edition, June 10, p. 38.

Kingdon, John
1984 Agendas, Alternatives, and Public Policies. Boston: Little, Brown, and Co.

Morone, James A., and Andrew B. Dunham
1985 "Slouching towards NHI: the new politics of health care." Yale Journal of Regulation 2.

Page, Benjamin I., Robert Y. Shapiro, and Glenn R. Dempsey
1987 "What moves public opinion," American Political Science Review 81, in press.

Payne, Stanley L.
1946 "Some opinion research principles developed through studies of social medicine." Public Opinion Quarterly 10: 93–98.

Pescosolido, Bernice A., Carol A. Boyer, and Wai Ying Tsui
1985 "Medical care in the welfare state: a cross-national study of public evaluations." Journal of Health and Social Behavior 26: 276–97.

Shapiro, Robert Y., and John M. Gillroy
1984 "The polls: regulation," Part I and Part II. Public Opinion Quarterly 48: 531–42, 666–77.

Shapiro, Robert Y., and Tom W. Smith
1985 "The polls: Social Security." Public Opinion Quarterly 49: 561–72.

Sidel, Victor W., and Ruth Sidel (eds.)
1984 Reforming Medicine: Lessons of the Last Quarter Century. New York: Pantheon Books.

Steiber, Steven R., and Leonard A. Ferber
1981 "Support for national health insurance: intercohort differentials." Public Opinion quarterly 45:179–98.

Taylor, Humphrey
1985 "Healing the health care system." Public Opinion, Aug/Sep 1985, pp. 16–20, 60.

Support for Government Action and Spending for Medical Care

NORC, Roper: I'm going to read off a number of different things. For each would you tell me whether you think it is a privilege that a person should have to earn, or a right to which he is entitled as a citizen? . . . Adequate medical care.

	NORC 4/1968[a]	*Roper 2/1975*	*Roper 2/1978*
Privilege	14%	21%	19%
Right	86	76	79
Don't know	1	3	2

[a] Next, I'd like you to tell me whether you agree strongly, agree somewhat, disagree somewhat, or disagree strongly with each of the following statements. . . . Every person has a right to receive free medical care if he needs it but cannot afford it.

NORC–GSS: In general, some people think that it is the responsibility of the government in Washington to see to it that people have help in paying for doctors and hospital bills. Others think that these matters are not the responsibility of the federal government and that people should take care of these things themselves. Where do you put yourself on this [6-point] scale or haven't you made up your mind on this?

	1975	*1983*	*1984*
1. I strongly agree it is the responsibility of government to help	36%	26%	24%
2.	13	19	18
3. I agree with both answers	28	32	34
4.	8	10	12
5. I strongly agree people should take care of themselves	13	10	8
6. DK/NA	2	4	4

ROPER: Turning now to the business of the country—we are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. . . . Improving and protecting the nation's health.

	10/71	*12/73*	*12/74*	*12/75*	*12/76*	*12/77*	*12/78*	*12/79*	*12/80*	*12/81*	*12/82*	*12/83*
Too much	4%	5%	5%	6%	7%	6%	7%	7%	6%	6%	5%	5%
Too little	55	57%	60%	57%	55%	58	60	59	58	56	56	56
About right	30	32	27	29	31	28	27	29	31	33	35	33
DK/NA	11	7	8	8	6	7	7	5	6	6	4	6

NORC–GSS: We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. . . . Are we spending too much, too little, or about the right amount on . . . improving and protecting the nation's health?

	1973	*1974*	*1975*	*1976*	*1977*	*1978*	*1980*	*1982*	*1983*	*1984*	*1985*
Too much	5%	5%	5%	5%	7%	7%	8%	6%	5%	7%	6%
About right	31	28	28	31	32	34	34	32	34	31	33
Too little	61	64	62	60	56	55	55	56	57	57	58
DK/NA	4	4	4	4	5	4	4	5	4	5	3
										n = 490	*n* = 751

CRI: In 1981, spending for health and health care in the United States totaled about 10% of the country's gross national product, which is the value of all the goods and services produced in the country. Compared to spending on things like food, housing, education, and national defense, do you think spending 10% of the gross national product on health care is too much, too little, or about the right amount?

	1982	*1983*	*1984*	*1985*[a]
Too much	28%	16%	19%	28%
Right amount	35	27	32	30
Too little	23	43	35	26
Don't know	14	14	14	16

[a] In 1985 spending . . . totaled about 11%. . . . Do you think spending 11%. . . .

TRENDEX. I would like to get your opinion on several areas of important government activities. As I read each one, please tell me if you would like to see the government do more, do less or do about the same amount as they have been on . . . Health measures.

	1/65[a]	*1/66*[a]	*1/68*	*4/68*	*10/68*	*1/69*	*7/69*	*4/70*	*7/70*	*4/71*
Do more	67%	59%	61%	72%	73%	68%	74%	78%	79%	78%
Do less	9	9	8	7	5	5	5	4	4	3
Do the same	23	31	29	19	20	26	19	16	15	16
DK/NA	1	—*	1	1	2	1	2	2	1	2

	10/71	*4/72*	*7/72*	*10/72*	*1/73*	*7/73*	*1/74*	*7/74*	*10/74*	*1/75*
Do more	82%	80%	81%	82%	74%	84%	76%	82%	76%	73%
Do less	3	4	4	3	7	3	4	4	6	7
Do the same	13	14	13	13	19	12	18	13	17	19
DK/NA	2	1	2	1	6	1	2	1	1	2

	4/75	*7/75*	*10/75*	*4/76*	*7/76*	*10/76*	*1/77*	*4/77*	*7/77*	*10/77*
Do more	73%	76%	70%	72%	75%	71%	70%	70%	73%	71%
Do less	7	7	11	9	8	10	8	10	8	9
Do the same	18	16	18	17	15	18	20	18	18	18
DK/NA	1	1	2	2	2	1	2	1	1	2

	1/78	*4/78*	*10/78*	*1/79*	*4/79*	*7/79*	*10/79*	*1/80*	*4/80*	*7/80*
Do more	69%	71%	69%	67%	70%	70%	70%	63%	65%	67%
Do less	10	11	11	14	10	9	11	14	13	11
Do the same	19	17	18	18	18	20	18	21	21	21
DK/NA	3	2	1	1	1	1	1	2	1	—*

	10/80	*1/81*	*4/81*	*7/81*	*10/81*	*1/82*	*4/82*	*7/82*	*10/82*
Do more	66%	60%	61%	66%	66%	60%	64%	70%	69%
Do less	12	14	13	8	10	13	8	9	10
Do the same	20	25	24	20	23	25	27	20	20
DK/NA	1	1	1	2	1	2	1	1	1

* Less than one 0.5 percent.

[a] I would like to get your opinion on several areas of important expenditures, first, on the part of the government. As I read each one, please tell me if you would like to see the government spend more, spend less or spend about the same amount as they have been on . . . health measures.

TRENDEX: I would like to get your opinion on several areas of important *government* activities. As I read each one, please tell me if you would like the government to do more, do less, or about the same as they have been . . . on health measures. I notice you said you would like the government to do more on (health measures). Would you favor this increased activity if it required an increase in taxes?

	4/72	*4/72*	*10/78*	*4/79*	*4/81*	*4/82*
Yes	61%	61%	69%	65%	62%	65%
No	30	29	26	26	29	28
DK/NA	9	11	6	9	8	6
	n = 872	*n* = 716	*n* = 708	*n* = 714	*n* = 633	*n* = 643

CRI: In the next two decades there will be a big increase in the number of older people in our society. Would you be willing to have your taxes increased to provide long-term care in nursing homes for those elderly people who need it, or not?

	1981	*1982*	*1983*	*1985*
Yes	49%	53%	43%	58%
No	31	27	32	25
Not sure	20	20	24	17

National Health Insurance and Health Care Reform

NBC: Do you favor or oppose a national program of health insurance giving medical and hospital care to everyone, with the government paying part of the cost?

	8/78	*11/78*	*7/79*
Favor	64%	61%	61%
Oppose	31	33	31
Not sure	5	6	7

SRC/CPS,MOR: There is much concern about rapid rise in medical and hospital costs. Some feel there should be a government health insurance plan which would cover all medical and hospital expenses. Others feel that medical expenses should be paid by individuals and through private insurance like Blue Cross. Where would you place yourself on this scale, or haven't you thought much about this?

	SRC/CPS		*MOR*	*MOR*	*SRC/CPS*	*SRC/CPS*	*MOR*	*SRC/CPS*
	1970	*1972*[a]	*1974*	*1976*	*1976*[a]	*1978*[a]	*1979*	*1984*[b]
1. Government insurance plan	29%	31%	32%	30%	28%	28%	24%	16%
2.	9	7	10	9	9	10	9	10
3.	8	8	7	9	7	8	9	11
4.	15	14	15	17	12	13	18	20
5.	6	7	7	9	9	9	10	16
6.	10	6	7	7	10	10	11	14
7. Private insurance plan	24	28	21	20	25	22	20	12

[a] Omits "health."
[b] Omits "health." . . . Blue Cross or other company-paid plans.

NYT/CBS: Do you favor or oppose . . . national health insurance, financed by tax money, and paying for most forms of health care?

	8–9/76[a]	*7/77*[b]	*2/80*	*3/80*	*4/81*[c]
Favor	49%	60%	50%	46%	52%
Oppose	39	33	41	43	37
No opinion	9	7	9	11	11

[a] Do you favor or oppose a national program of health insurance with the government paying the cost?
[b] Now, we'd like to ask you about some of the things the government does with our taxes. Do you approve or disapprove . . . of a national health care program for all people—with taxes paying the cost?
[c] Do you favor or oppose national health insurance, which would be financed by tax money, paying for most forms of health care?

ROPER: There are differing opinions as to how health care insurance should be provided. Some people favor the present system of private insurance and health plans. Others say we should have a national system of health insurance provided through the government and paid for like Social Security now is. Which do you favor—our present system of private health plans, or a national health insurance plan provided through the government?

	9/73	*2/74*	*2/75*	*9/75*	*10/76*	*9/77*	*9/78*	*2/79*	*9/83*
Present private plan	50%	43%	40%	50%	52%	45%	50%	40%	49%
National plan	40	45	46	41	39	47	40	44	41
Don't know	10	12	14	9	10	9	10	15	10

ROPER: There are many problems facing our nation today. But at certain times some things are more important than others, and need more attention from our Federal government than others. I'd like to know for each of the things on this list whether you think it is something the government should be making a *major* effort on now, or something the government should be making *some* effort on now, or something not needing any particular government effort now . . . Government effort to set up a national health insurance system.

	6/79	*6/80*	*6/81*	*6/82*	*6/83*	*6/84*
Major effort	50%	41%	45%	44%	50%	53%
Some effort	26	31	28	30	28	27
Needs no effort	18	23	21	21	18	15
Don't know	6	5	5	6	4	4

YSW/HIAA: Attitudes toward national health insurance with tax increase.

	1978	*1979*	*1980*	*1982*
Favor strongly	16%	21%	19%	16%
Favor somewhat	21	17	22	18
Probably favor	28	20	22	25
Probably oppose	14	15	14	15
Oppose somewhat	9	11	9	10
Oppose strongly	11	17	15	16
DK/NA	1	—	—	—

YSW/HIAA: Private health insurance companies give better service to the people they insure than a government health insurance program would.

	1977	*1979*	*1980*	*1982*
Agree	33%	34%	33%	29%
No strong opinion	46	43	46	50
Disagree	18	20	20	18

YSW/HIAA: The competition among the various health insurance companies leads to benefits for the consumer that would not be available through a government-run health insurance program.

	1977	*1979*	*1980*	*1982*
Agree	37%	30%	30%	30%
No strong opinion	45	48	50	50
Disagree	16	19	19	17

YSW/HIAA: National health insurance: Catastrophic vs. Comprehensive

	1979	*1980*	*1982*
No new government health care program should be enacted.	25%	21%	25%
A new government health care program should be enacted that would cover people only for those medical expenses that could ruin them financially.	33	42	37
A new government health care program should be enacted that would cover almost all the health care cost people have.	39	35	37

HARRIS/KFF: Which of the following statements comes closest in expressing your overall view on the American health care system?

	1982	*1983*	*1984*
On the whole, the health care system works pretty well, only minor changes are necessary to make it work better.	19%	21%	26%
There are some good things in our health care system, but fundamental changes are needed to make it work better.	47	50	49
The American health care system has so much wrong with it that we need to completely rebuild it.	28	25	21

Government Regulation

KPA: If there were less government regulation of health care, in your opinion, what would probably happen to the quality of health care?

	1979	*1981*
Improve a great deal	5%	3%
Improve somewhat	13	19
No change	25	20
Decline a great deal	11	10
Decline somewhat	29	33
DK	17	15

KPA/AMA: Government regulation of health care has some good effects and some bad effects. Do you feel that the good outweighs the bad, or do you feel that the bad outweighs the good?

	1979	*1981*
Good outweighs bad	51%	50%
Bad outweighs good	25	29
DK	24	21

CRI: For each specific industry I mention, please tell me whether you think there is now too much or too little government regulation . . . Health care industry.

	1982	*1983*	*1985*
Too much	17%	19%	22%
Right amount	19	16	18
Too little	53	54	49
DK	12	12	12

Medical Care Costs

ROPER: How do you feel about the cost of the medical care you receive—do you feel the cost is very reasonable, fairly reasonable, somewhat unreasonable, or very unreasonable?

	10/73	*10/75*	*10/76*	*10/78*
Very reasonable	12%	13%	12%	10%
Fairly reasonable	30	28	32	28
Somewhat unreasonable	28	30	27	28
Very unreasonable	27	26	27	31
DK	3	3	2	3

YSW/HIAA: The price of health insurance is getting too high for the average family to afford.

	1977	*1978*	*1979*	*1982*
Agree	67%	70%	72%	71%
No strong opinion	24	21	20	23
Disagree	8	8	7	5

KPA/AMA: Now thinking about everything that has to do with health care, medicine, and doctors, what do you feel is the main problem facing health care and medicine in the United States today?

	1981	*1982*	*1983*	*1984*
Cost	55%	62%	65%	68%
Quality	14	13	14	12
Access	6	8	6	5
Other	25	17	15	15

SRC/CPS, NYT/CBS: The government ought to help people to get doctors and hospital care at low cost.

	SRC/CPS *1960*[a]	*SRC/CPS* *1962*[b]	*NYT/CBS* *1978*
Agree	75%	72%	85%
Disagree	25	18	15

[a] The government ought to help people get doctors and hospital care at low cost. Agree strongly, agree but not very strongly . . . disagree but not very strongly, disagree strongly.

[b] The government ought to help people get doctors and hospital care at low cost. Do you have an opinion on this or not. (If yes) Do you agree the government should do this or do you think the government should not do it. Yes, Yes—qualified, for the aged (Medicare) or for those financially unable. No, no—qualified, except for aged (Medicare).

CRI: I'm going to read you some particular measures that have been proposed to help control health care costs. I'd like you to tell me whether you would strongly favor, somewhat favor, somewhat oppose, or strongly oppose each one. . . . Government price controls and regulations on what doctors and hospitals can charge.

	1979	*1981*	*1982*	*1984*
Strongly favor	38%	36%	41%	50%
Somewhat favor	30	25	26	27
Somewhat oppose	15	14	13	9
Strongly oppose	12	19	16	10
DK	4	6	4	4

. . . Closing inefficient hospitals or merging them into more efficient units when they are in the same vicinity.

	1979	*1981*	*1982*	*1984*
Strongly favor	56%	48%	52%	51%
Somewhat favor	29	26	28	27
Somewhat oppose	7	9	7	9
Strongly oppose	3	8	7	5
DK	5	8	7	8

. . . Giving insurance companies more power to negotiate fees with doctors and hospitals.

	1979	*1981*	*1982*	*1984*
Strongly favor	33%	28%	33%	35%
Somewhat favor	33	29	30	30
Somewhat oppose	12	12	10	11
Strongly oppose	12	18	16	12
DK	10	13	10	11

. . . Stimulating competition among hospitals and doctors by encouraging consumers to compare costs and shop around for medical care.

	1981	*1982*
Strongly favor	43%	49%
Somewhat favor	31	30
Somewhat oppose	9	7
Strongly oppose	10	8
DK	7	6

Attitudes Toward Health Industry

ROPER: Thinking of all your medical care needs, not just your family doctor, generally speaking, how satisfied would you say you are with the *quality* of the medical care you get—very satisfied, fairly well satisfied, not too satisfied, or not at all satisfied?

	10/73	*10/75*	*10/76*	*9/78*	*9/83*
Very satisfied	49%	54%	51%	45%	48%
Fairly well satisfied	34	33	34	38	35
Not too satisfied	12	8	9	10	9
Not at all satisfied	4	4	4	5	5
DK/NA	1	2	2	2	2

ROPER: And how satisfied are you with the availability of medical care when you need it—very satisfied, fairly well satisfied, not too satisfied, or not at all satisfied?

	10/73	*10/75*	*10/76*	*9/78*	*9/83*
Very satisfied	47%	52%	46%	46%	51%
Fairly well satisfied	28	29	32	36	33
Not too satisfied	15	9	11	10	8
Not at all satisfied	6	5	7	5	5
DK/NA	4	5	4	3	3

CRI: Now, I'm going to read you a list of various institutions or types of industries. After each one, I'd like you to tell me whether you have a very favorable, somewhat favorable, somewhat unfavorable, or very unfavorable opinion of these industries . . . The health care industry, including doctors and hospitals.

	1979	*1980*	*1981*	*1982*	*1983*	*1984*
Very favorable	9%	14%	14%	25%	14%	19%
Favorable	37	38	41	34	39	40
Unfavorable	30	25	25	21	24	24
Very unfavorable	21	21	17	18	20	15
DK	3	3	3	2	2	2

KPA/AMA: Please tell me if you agree with each of the following statements about doctors.

	(Percentage Agree)		
	1982	*1983*	*1984*
Most physicians are accessible in an emergency	NA	56%	53%
Most physicians are available for routine health care problems	NA	69	NA
Doctors are usually up to date on the latest advances in medicine	71	72	71
Most doctors are genuinely dedicated to helping people	80	76	NA
Most doctors take a genuine interest in their patients	68	62	62
Doctors act like they're better than other people	33	35	38
Doctors fees are usually reasonable	42	32	27
Doctors are too interested in making money	60	66	67
Doctors are active in trying to hold down the cost of medical care	NA	22	21
Most doctors spend enough time with their patients	46	38	NA

	(Percentage Agree)		
	1982	*1983*	*1984*
Doctors usually explain things well to their patients	55	49	44
Poor people are able to get needed medical care	48	41	43
The elderly are able to get needed medical care	52	50	45
People are beginning to lose faith in doctors	62	66	68

YSW/HIAA: Degree of satisfaction with present health insurance coverage.

	1975	*1976*	*1977*	*1978*	*1982*
Very satisfied	44%	42%	46%	40%	40%
Somewhat satisfied	36	37	36	36	41
Somewhat dissatisfied	9	11	9	13	8
Very dissatisfied	7	7	6	7	5
DK/NA	4	3	3	4	6

Correction

In "The Polls: Environmental Protection" (*POQ* 50 (2):273) the following figures are the correct ones: NORC-GSS, 1985, Too much: 8%, Dk: 4%; Roper, 1983, Too far: 14%.

Public Opinion Quarterly

"Some of my best friends are black . . .": Interracial Friendship and Whites' Racial Attitudes — *Mary R. Jackman and Marie Crane*

Yuppies — *John L. Hammond*

Do Yuppies Matter? Competing Explanations of Their Political Distinctiveness — *Michael Delli Carpini and Lee Sigelman*

The Perceived Threat of Nuclear War, Salience, and Open Questions — *Howard Schuman, Jacob Ludwig, and Jon A. Krosnick*

Ecological Concern, Attitudes, and Social Norms in Voting Behavior — *James D. Gill, Lawrence A. Crosby, and James R. Taylor*

Change and Stability in Presidential Popularity at the Individual Level — *Kent L. Tedin*

Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos — *Stephen D. Reese, Wayne A. Danielson, Pamela J. Shoemaker, Tsan-Kuo Chang, and Huei-Ling Hsu*

THE POLLS: The Most Admired Man and Woman — *Tom W. Smith*

BOOK REVIEWS • BOOK NOTES

INDEX TO VOLUME 50

WINTER 1986

Published for The American Association for Public Opinion Research
by The University of Chicago Press

Public Opinion Quarterly

JOURNAL OF THE AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

AAPOR ADVISORY COMMITTEE

LEO BOGART BURNS W. ROPER BERNARD ROSHCO
ELEANOR SINGER, EX OFFICIO
HERBERT ABELSON AND NORMAN M. BRADBURN, COUNCILLORS-AT-LARGE

Editor: ELEANOR SINGER *Managing Editor:* INA C. JAEGER
Department Editors:
PHILIP A. HARDING, *News and Notes*
MARC B. GLASSMAN, *Book Reviews*
TOM W. SMITH, *The Polls*

EDITORIAL BOARD

JAMES R. BENIGER
JOAN S. BLACK
NORMAN M. BRADBURN
CHARLES F. CANNELL
MURIEL G. CANTOR
ALBERT H. CANTRIL
DONALD DILLMAN
MARTIN R. FRANKEL
HERBERT J. GANS
ALBERT E. GOLLIN
ROBERT HORNIK
MARY R. JACKMAN
F. THOMAS JUSTER
RONALD C. KESSLER
MARK R. LEVY
PHILIP MEYER
J. RONALD MILAVSKY
WARREN J. MITOFSKY
ELISABETH NOELLE-NEUMANN
STANLEY PRESSER
DONALD F. ROBERTS
MORRIS ROSENBERG
CHARLES R. WRIGHT
DANIEL YANKELOVICH

Public Opinion Quarterly (ISSN 0033-362X) Volume #50, 4 issues (1986)

© 1986 by the American Association for Public Opinion Research.

Published quarterly: Spring, Summer, Autumn, and Winter, by The University of Chicago Press, Chicago. Second-class postage paid at Chicago, IL. Subscription rates, U.S.A.: institutions 1 year $38.00; individuals 1 year $19.00. AAPOR member subscription rate, U.S.A.: individuals 1 year $17.00. Student subscription rate (with copy of validated ID), U.S.A.: individuals 1 year $17.00. Other countries add $3.00 for each year's subscription to cover postage.

Japanese subscription agent: Kinokuniya Company Ltd.

POSTMASTER: Please send all change of address notices to Public Opinion Quarterly, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Mail orders and direct subscription information to: Journals Fulfillment Department, The University of Chicago Press, P.O. Box 37005, Chicago, IL 60637.

Single copy and back volume information furnished upon written request. Claims for missing issues can be honored only up to three months from date of issue. Indexed by Public Affairs Information Services, Social Sciences Index, United States Political Science Documents, Combined Retrospective Index Sets (CRIS), and Political Science Abstracts (PSA). Available from University Microfilms International (UMI).

Editorial Office: Journalism Building, Columbia University
116th Street and Broadway, New York, New York 10027
Business, Subscriptions, Back Issues: The University of Chicago Press,
5801 South Ellis, Chicago, IL 60637
AAPOR: Post Office Box 17, Princeton, New Jersey 08540
Subscribers and members of AAPOR are urged to notify the Subscription Office as soon as possible of changes of address. In doing so, give the old as well as the new address.

PUBLISHED BY THE UNIVERSITY OF CHICAGO PRESS

Public Opinion Quarterly

VOLUME 50 WINTER 1986 NUMBER 4

"Some of my best friends are black . . .": Interracial Friendship and Whites' Racial Attitudes
Mary R. Jackman and Marie Crane 459

Yuppies *John L. Hammond* 487

Do Yuppies Matter? Competing Explanations of Their Political Distinctiveness *Michael Delli Carpini and Lee Sigelman* 502

The Perceived Threat of Nuclear War, Salience, and Open Questions
Howard Schuman, Jacob Ludwig, and Jon A. Krosnick 519

Ecological Concern, Attitudes, and Social Norms in Voting Behavior
James D. Gill, Lawrence A. Crosby, and James R. Taylor 537

Change and Stability in Presidential Popularity at the Individual Level *Kent L. Tedin* 555

Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos *Stephen D. Reese, Wayne A. Danielson, Pamela J. Shoemaker, Tsan-Kuo Chang, and Huei-Ling Hsu* 563

THE POLLS: The Most Admired Man and Woman
Tom W. Smith 573

BOOK REVIEWS *Marc B. Glassman, Editor* 584

Howard Schuman, Charlotte Steeh, and Laurence Bobo, *Racial Attitudes in America: Trends and Interpretations*
Norval D. Glenn 584

Richard F. Hamilton and James D. Wright, *The State of the Masses* *Norman R. Luttbeg* 586

BOOK NOTES *Phyllis M. Endreny, Guest Editor* 588

INDEX TO VOLUME 50 592

NOTICE TO CONTRIBUTORS

Manuscripts to be considered for publication should be sent to: Dr. Howard Schuman, Editor, Public Opinion Quarterly, Survey Research Center, Institute for Social Research, P.O. Box 1248, University of Michigan, Ann Arbor, Michigan 48106.

Four copies of each article should be submitted. Authors are urged not to submit original typescripts. Copies of manuscripts which are reviewed become the property of the journal, and unused copies will be returned only if the author provides a stamped, self-addressed envelope. Bibliographic-references, footnotes, and tables should be typed to conform to the printed style of this journal.

All submissions that make use of survey data not archived at the Inter-University Consortium for Political and Social Research or the Roper Center must contain the following information: (1) A definition of the population from which the sample was drawn and a statement of the selection procedure sufficient to permit at least approximate replication. (2) The response rate and details of its calculation. (If response rate is not appropriate, the number of refusals.) (3) Dates the survey was conducted. (4) Exact wording of all questions used in the analysis.

Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright in the article to the publisher. This transfer will insure the widest possible dissemination of information under the U.S. Copyright Law.

PUBLIC OPINION QUARTERLY is hospitable to all points of view, provided that the material presented will help to illuminate problems of communication and public opinion. The American Association for Public Opinion Research assumes responsibility only for material signed by an officer of the Association in his or her official capacity. Statements of fact and opinion are made on the responsibility of the authors alone and do not represent the views of the editors.

PUBLIC OPINION QUARTERLY has been registered with the Copyright Clearance Center, Inc. Consent is given for copying of articles for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Center the per-copy fee stated in the code on the first page of each article for copying beyond that permitted by the U.S. Copyright Law. If no code appears on an article, the author has not given broad consent to copy and permission to copy must be obtained directly from the author. All articles published prior to 1978 may be copied for a per-copy fee of $1.00, also payable through the Center. The consent does not extend to other kinds of copying, such as for general distribution, resale, advertising and promotional purposes, or for creating new collective works.

"Some of my best friends are black . . .": Interracial Friendship and Whites' Racial Attitudes

MARY R. JACKMAN AND MARIE CRANE

See that man over there?
Yes.
Well, I hate him.
But you don't know him.
That's why I hate him.
—Parable (Allport, 1958:253)

THE NOTION that hostility has its basis in ignorance is a truism as much as the idea that social intimacy implies acceptance and equality. In discussions of whites' racial attitudes, these truisms are invoked re-

---

Abstract We examine the major tenets and assumptions of the well-known contact theory of prejudice, and we compare them with the more cynical reasoning implied by the infamous "Some of my best friends are black, but..." expression. After assessing the extant evidence for the contact theory, we use a unique set of national survey data to address the central postulates of that theory. We examine the racial beliefs, feelings, social dispositions, and policy views of whites who have contact with blacks as friends, acquaintances, or neighbors. Our results suggest that personal interracial contact is selective in its effects on whites' racial attitudes, that intimacy is less important than variety of contacts, and that any effects are contingent on the relative socioeconomic status of black contacts. On the basis of our analysis, we reassess the contact theory and propose a more political conception of the attitudes of dominant groups toward subordinates. We argue that the message contained in the relationship between personal contact with subordinates and intergroup attitudes is less benign than is suggested by the contact theory.

Mary R. Jackman is Professor of Sociology, the University of Michigan. During 1986–87 she is a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California. Marie Crane is Assistant Professor of Sociology, University of Texas at Austin. This research was supported by grants from NSF (SOC-78-16857) and NIMH (MH-26433), and by an NIMH Research Scientist Development Award to Mary Jackman (MH-00252). The authors are grateful to several people for their helpful comments on an earlier draft of this article: Barbara Anderson, Joe Feagin, Lowell Hargens, Robert Jackman, William Kelly, Donald Kinder, Barbara Reskin, Howard Schuman, James Sidanius, Brian Silver, Eleanor Singer, Walter Stephan, Blake Turner, and the anonymous reviewers.

Public Opinion Quarterly Vol. 50:459–486 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033–362X/86/0050–459/$2.50

peatedly. Physical separation of blacks from whites is such a marked feature of race relations in the United States that, indeed, only a tiny minority of whites could rightly claim that "some of their best friends" are black. The rarity of interracial friendship combined with the prevailing conservatism of whites' racial attitudes would thus seem to underscore the validity of Allport's parable. Yet there is alternative folk wisdom on the subject that seems equally plausible. The infamous "some of my best friends are black. . ." expression in the title of our article is much less sanguine about the effects of personal intimacy with blacks on interracial attitudes.

Friends make up the social fabric of our day-to-day lives. Hence the significance typically attached to the pattern of personal ties in the reinforcement or muting of subjective group boundaries (see both Marxist and pluralist accounts, e.g., Marx, 1964: 178–202; de Tocqueville, 1969; Polsby, 1980; Hodge and Treiman, 1968). In the analysis of whites' racial attitudes in the United States during the past few decades, the issue of personal intergroup contacts has seemed especially pertinent. Researchers in the 1940s and 1950s grappled with the widespread problem of white prejudice and discrimination in the historical context of a society marked by pervasive racial segregation. Motivated by the spirit of social reform, analysts sought to identify ways in which whites' personal contact with individual blacks might be manipulated by public policy to dissipate their prejudice toward blacks as a group. Out of this concern emerged one of the most prominent theories in the prejudice literature, known as the *contact theory*.

In this article, we outline the major tenets and assumptions of the contact theory, compare them with the ideas implicit in the "Some of my best friends are black, but..." expression, and briefly assess the evidence for the contact theory that was generated by some landmark experiments in social engineering. We then employ a unique set of national survey data to address a series of empirical questions dictated by the contact theory. On the basis of that analysis, we reassess the contact theory and propose a more political conception of the attitudes of dominant groups toward subordinates. We suggest that the relationship between personal contact with subordinates and intergroup attitudes contains a less benign message than that put forth by the contact theory.

The Contact Theory. This theory holds that the sharp rupture between the social lives of whites and blacks promotes whites' ignorance about blacks. This ignorance feeds erroneous, oversimplified, negative beliefs about blacks, which in turn engender feelings of hostility and discriminatory social and political predispositions toward blacks. By bringing whites into personal contact with blacks, erroneous images of blacks can be corrected and hostile dispositions softened (Myrdal, 1944; Deutsch and Collins, 1951, 1965; Wilner et al., 1952, 1955; Allport, 1958; Amir, 1969).

To be effective, the interracial contact must meet a specified set of conditions. First, the contact should not take place within a competitive context. Second, the contact must be sustained rather than episodic. Third, the contact must be personal, informal, and one-to-one. Fourth, the contact should have the approval of any relevant authorities. Finally, the setting in which the contact occurs must confer equal status on both parties rather than duplicate the racial status differential. Much interracial contact does not meet these conditions. Consider, for example, the contact between white mistress and black maid, or between white and black neighbors who pass each other daily on the street without personal interaction, or between black and white children who attend the same school by virtue of a school-busing program opposed by the local school board. These forms of contact are considered insufficient to remove whites' blinders and allow them to perceive blacks in a fresh light. In contrast, the contact that occurs between intimate, personal friends appears to meet optimally the con~~ditions of~~ the contact theory.

While the contact theory developed primarily out of a policy-oriented concern with proposals to reduce prejudice, its central tenets rest on important assumptions about the very nature of intergroup attitudes. Most fundamentally, proponents of the contact theory, like students of interracial attitudes generally, have assumed that intergroup attitudes express a parochial negativism, rather than political interests. This assumption is reflected in the concept of prejudice, which has been routinely used to define the problem. Witness Allport's classic definition of prejudice: "Ethnic prejudice is an *antipathy based upon a faulty and inflexible generalization.* It may be felt or expressed. It may be directed toward a group as a whole, or toward an individual because he is a member of that group" (Allport, 1958:10, emphasis added). Thus, negative intergroup attitudes are prejudiced attitudes that have an irrational basis and are permeated by feelings of hostility. This fundamental assumption has three important corollaries.

First, intergroup attitudes are interpreted primarily as a property of individuals. Researchers were drawn by individual-level variation in attitudes toward blacks and attempted to account for that variation by examining individual differences in personality, socialization, or interracial experience. The contact theory focused on the last as a potential policy tool.[1] Second, if negative intergroup attitudes are founded in irrationality and misinformation, the way to positive attitudes is with rationality and correct information. Serious differences do not exist between blacks and whites, and thus exposure to blacks (under the right conditions) will reveal the falsity of negative beliefs about blacks. Third, because discriminatory behavior predispositions toward blacks reflect a feeling of antipathy, the way to nondiscriminatory predispositions is to generate positive feelings toward blacks. In short, the locus of the problem was in the individual, in erroneous generalizations, and in the connection between personal feelings of antipathy and discriminatory predispositions. These all pointed logically to the probable efficacy of a policy that would foster close personal friendships between individual whites and blacks.

---

[1] Bound by a similar conception of the problem and by a common social problems motivation, we explored alternative theories as mutually compatible solutions.

"Some of my best friends are black, but..." We are all familiar with this infamous expression, which suggests a more insidious relationship between personal friendship and intergroup attitudes. First, it implies that I separate my personal relationship with an individual who is black from my impersonal relationship with the social category "blacks." Second, because I have a personal connection with blacks, I feel I am an authority about blacks and therefore I am confident in my judgments. Third, I feel that having one or two black friends testifies that I am personally unbiased and fair in my assessments of blacks, which again massages my complacence about my racial views. In short, having one or two black friends gives me a license to think what I please about the group as a whole.

The ideas implied by this common expression have not found their way into an alternative theory about the relation between personal friendship and intergroup attitudes. The notion that something powerful blocks the benign generalization from one's friends to other people suggests, however, that intergroup attitudes may be driven by a force that is not susceptible to social reform.

Evidence for the Contact Theory from the Housing Project Studies. Several landmark social experiments in racial integration were conducted in the 1940s and 1950s, most notably in the wartime army (Stouffer, 1949), on merchant ships (Brophy, 1945), and in low-income public housing projects (Deutsch and Collins, 1951, 1965; Wilner et al., 1952, 1955). Each study reported positive effects on whites' racial attitudes, although the generality of the attitudes to broader social situations was sometimes questionable. By far the most comprehensive and ingenious studies were those conducted by Deutsch and Collins, and Wilner et al. These analysts compared segregated housing projects with projects that had been integrated in a variety of ways. And they used attitude measures that reflected racial beliefs, feelings, and social predispositions—in relation to both blacks living in the project and blacks in general.

In their analysis of the behavior and attitudes of the white housewives living in these varying conditions, these investigators found that proximity had a pronounced positive effect on the level of informal interaction and friendship with blacks, and that both of these factors in turn had a strong positive effect on attitudes toward blacks. The housing studies stressed the importance of sustained, intimate contacts in order to provide a compelling perceptual stimulus and a strong motivational basis for attitude change. Wilner et al. observed that "the more intimate the contact, the more favorable the attitude—without exception" (1955:99). In a subsequent essay, Cook reflected: "One of the clearest findings of studies on the relation between intergroup contact and attitude change is that, while individuals rather quickly come to accept and even approve of association with members of another social group in situations of the type where they have experienced such association, this approval is not likely to be generalized to other situations unless the individuals have quite close personal relationships with members of the other group" (1963:41–42).

The public-housing studies left two important empirical questions unanswered. First, are the strong positive outcomes observed in the relatively controlled, bounded, and stable setting of the public housing projects generalizable to contact situations in ordinary (unplanned) neighborhoods and workplaces? There has been a paucity of cross-sectional survey data addressing the empirical questions posed by the contact theory.[2] Second, we still have little idea of the relationship between personal interracial contact and racial predispositions of a more political nature. The housing-project studies were understandably absorbed with the more immediate question of whether whites could learn to accept and respect blacks in personal terms. They measured racial beliefs, feelings, and social dispositions, but they did not measure support for governmental intervention to promote black civil rights and racial equality. Do the positive effects observed for the more personal, affectively loaded measures also hold for more political orientations toward blacks? Is the issue of equality tied to the issue of personal acceptance?

Analysis

DATA AND EMPIRICAL QUESTIONS

The data for this article come from a national probability survey of adults aged 18 and over residing in the 48 contiguous United States.

---

[2] Two interesting, and inconclusive, studies of the effects of neighborhood integration in specific communities are reported in Williams (1964) and Hamilton and Bishop (1976).

The survey was administered in the fall of 1975 by the Survey Research Center of the Institute for Social Research at The University of Michigan. The response rate was just under 70 percent, and face-to-face interviews were conducted with 1914 respondents, of whom 1648 were whites. The data contain two sets of measures that are uniquely suited to the analysis of personal interracial contact and white racial attitudes.

First, there are unusually detailed and complete data on the *racial composition of people's friendship circles*. Questions about respondents' friends were introduced as follows:

> I would like to ask you some questions about the people you consider your good friends—by good friends I mean adults you enjoy getting together with at least once a month or so and any other adults who live elsewhere that you try to keep in close touch with by calling or writing.

Respondents were then handed a "Friends Sheet" on which they listed the first name of each of their "good friends," followed by a variety of other information including each friend's race and occupation.

These data are considerably more detailed and specific than previous attempts in surveys to assess the extent of interracial friendship, where the standard question has been of the global type, "Are your friends all white, mostly white, about half white and half black, mostly black, or all black?" (see, e.g., the SRC Election Studies). At the same time, our data give more complete information on the respondent's entire circle of friends than does the standard procedure in the literature on friendship networks of seeking individual-level data on the respondent's three best friends only.

After completing the Friends Sheet, respondents were asked, "Are there any other people you keep in touch with or get together with occasionally?" If they said "yes," they were asked, "Which of the categories on this card best describes how many blacks and whites are among these people? [All black; Mostly black; About half and half; Mostly white; All white]." This global question allows us to reach beyond the bounds of good friendship and make an approximate assessment of the *racial composition of people's acquaintance circles* as well.

Those data are supplemented with information on the *racial composition of the respondent's neighborhood and place of work,* about any *informal interaction with blacks* in each of those settings ("Do you and any of your black neighbors ever do anything like talk informally over coffee or in the street, lend each other things, or help each other out with baby-sitting or chores?"; "Do you and any of the blacks at work ever have lunch together or talk informally over coffee about things not related to work?''; if ''yes'': ''How often do you do this—very often or not too often?'' [some respondents volunteered ''sometimes'']), and about *exposure to blacks in previous neighborhoods and workplaces* (e.g., ''Have you ever lived in another neighborhood that had one or more black residents? [Yes/No]'').

The second set of measures is an unusually comprehensive series of questions to reflect whites' *racial attitudes*. The items appear in Appendix A. They encompass all the primary elements of ''prejudice'': beliefs about personality trait differences between blacks and whites, feelings of emotional preference for whites over blacks, personal predispositions for contact with or avoidance of blacks, and policy orientations relating to government intervention on behalf of blacks. All of the attitude items have neutrally balanced response-options in order to avoid response biases, and the most sensitive items (on racial beliefs and feelings) were contained in self-administered booklets in order to minimize social desirability pressures.

Our analysis begins with the preliminary issue of how some whites come to acquire black friends, by choice or by chance. Next, we address the baseline question of the relationship between interracial friendship and attitudes toward blacks, and we compare the effect of friendships and more casual acquaintances with blacks. Do people seem to generalize from personal contacts to the group as a whole? Are the close personal ties of friendship more compelling or are the racial attitudes of whites as likely to be influenced by the looser (and less rarely acquired) ties of acquaintanceship with blacks? Third, we compare the effects of such personal interracial contact with the effects of sheer proximity to blacks. Are the conditions specified by the contact theory necessary for a change in attitudes, or is mere exposure to blacks sufficient to break down prejudiced attitudes in this segregated society? As we address these questions, we also examine the evidence for a ''tokenism'' effect. Is one black friend or acquaintance enough to change a white's racial attitudes, or is such personally observed data disregarded unless there is a less token presence of blacks in a white's social environment? Finally, we abandon the idea that our friends are all equals in our affectionate eyes, and examine the racial attitudes of whites who have black friends, according to the relative socioeconomic status of their black and white friends. For each of these issues, we can compare the effects of interracial contact on whites' interracial beliefs, feelings, social predispositions, and policy orientations to assess the extent to which they respond in a unitary or divergent way to the establishment of personal ties with blacks.[3]

---

[3] In addressing these questions, we rely on percentage cross-tabulations, for several reasons. First, we wish to reduce the data as little as possible. Second, our purpose is not

ACQUIRING A BLACK FRIEND

Consideration of the relationship between personal interracial contact and racial attitudes raises the possibility of self-selection: we like to think that something as personal as our circle of good friends reflects our own choices and not factors that are out of our control, and we know that many whites would prefer to avoid or minimize their contact with blacks. In this article, however, we assume that whites' initial racial attitudes have little to do with their likelihood of acquiring a black friend, and that it is therefore appropriate to ask whether friendship with individual blacks influences the way whites think about blacks as a group. There may be occasional extreme cases of whites who either eagerly seek out or steadfastly avoid contact with blacks, but for the vast majority, other exigencies overwhelm racial considerations in the selection of friends. Much as many whites might prefer to avoid blacks, economic and practical concerns constrain individuals to take a job regardless of blacks' presence in the workplace, to reside in a neighborhood regardless of blacks' presence there, and so on.[4] Thus hemmed in, the exercise of personal choice is itself unlikely to be a singular response to race, but instead is a response to the configuration of personal characteristics that each individual contact offers (e.g., interests, lifestyle, personality).

The validity of our argument cannot be demonstrated unequivocally with cross-sectional data, but its assumptions and empirical expectations differ sharply enough from those of the self-selection thesis to present several critical issues that can be investigated empirically. The self-selection thesis assumes that whites' racial attitudes determine whether they enter or avoid situations where blacks are present, or, if forced into an interracial situation, whether they engage in personal contact with blacks. Our argument assumes that the racial composition of individuals' friendship circles primarily reflects the availability of blacks and whites in their day-to-day lives. A series of empirical expec-

---

to account for the overall variance in whites' racial attitudes, which is the kind of endeavor for which a more formal statistical procedure (such as regression analysis) is designed. Instead, we want to focus on the contact theory's more modest prediction that when personal contact with blacks takes place, prejudice is reduced. Finally, given the relatively small number of whites who have personal contact with blacks, a more formal statistical procedure would provide a less sensitive indication of any contact effects.

[4] This argument appears to fly in the face of the much-publicized phenomenon of "white flight." However, the empirical evidence for white flight from racially integrated schools is mixed (see, e.g., Farley et al., 1980). Besides, to the extent that white flight from blacks does occur in neighborhoods and schools, it is a collective rather than an individual phenomenon: individual decisions to leave are influenced by a variety of factors, including the behavior of others, assessments about future property values, school quality, and so on.

tations derives from these assumptions, which can be subsumed under two broad issues: (1) the relationship between proximity to blacks and personal contact with blacks, and (2) the nature of the association between interracial contact and racial attitudes.

Proximity and Friendship. The literature on sociometric choice has taught us that we make friends with people who are most readily available to us (Festinger et al., 1950; Berscheid and Walster, 1969; Vander Zanden, 1984). While friends reflect an element of personal choice, they do not reflect a free choice: we are most likely to become friendly with those who are thrown consistently in our path. In their classic study of friendship patterns in a student housing project, Festinger et al. (1950) showed that proximity was the single most important factor in determining those patterns: "It is a fair summary to say that the two major factors affecting the friendships which developed were (1) sheer distance between houses and (2) the direction in which a house faced" (Festinger, 1951:156). Their discovery was both profound and obvious. Our lives are sufficiently constrained by our work and our families that the path of least resistance is to find our friends from among those who find us. With such people we share enough common social turf to have an uncontrived entrée to their acquaintance and repeated opportunities for casual follow-up. The process is so pervasive that it is obvious to recognize at the same time as it is imperceptible to us as we conduct our social lives.

If interracial friendships are governed by the same principles, personal contact with blacks should increase with whites' proximity to blacks. On the other hand, if racial preference intervenes in the process, and whites' attitudes determine whether they seek, accept, or avoid contact with blacks, the relative proximity of blacks should not be related to the amount of contact that whites have with blacks.[5] This issue was addressed by both Deutsch and Collins (1951) and Wilner et al. (1952), who reported a strong relationship between the degree of proximity to blacks and the extent of informal interaction and friendship with blacks. More recent work on classroom racial composition and cross-race friendship choice also demonstrated the generality of the proximity principle (Hallinan, 1982; Hallinan and Smith, 1985). Our own data also indicate that whites' proximity to blacks in ordinary neighborhoods and workplaces is strongly related to their likelihood and frequency of personally interacting with blacks, as well as to their likelihood of having black friends and acquaintances. Those relationships are reported in the first part of Appendix B.

---

[5] If the proximity results from purposive behavior on the part of whites, interactions with blacks should be highly probable among all whites with some proximity to blacks, regardless of the degree of proximity. If it is the degree of personal contact that reflects racially motivated purposive behavior within interracial situations, we would expect the probability of interaction with blacks to be equally high or low regardless of the relative proportion of blacks in an integrated situation.

Interracial Contact and Racial Attitudes. The self-selection thesis assumes a strong and unequivocal association between interracial contact and racial attitudes. In contrast, our assumption about the process of interracial contact leaves its association with attitudes an open question. The results of the analyses reported in this article indicate that neither the strength of the relationship between interracial contact and racial attitudes nor its specific nuances conform to the logical expectations of the self-selection thesis. It would be premature to discuss that evidence at this point: relevant data from Tables 1, 2, 3 are identified in the second part of Appendix B.

THE INFLUENCE OF BLACK FRIENDS AND ACQUAINTANCES

We now examine the significance for whites' racial attitudes of the establishment of personal ties to blacks. Are whites who have one or more black friends different in their racial attitudes from whites who have no black friends? And what are the attitudinal effects of the looser, more easily acquired ties of casual acquaintance with blacks?

Only 9.4 percent of whites can name a "good friend" who is black, while 21.4 percent have at least one black acquaintance. Because interracial friendships and acquaintances are both fostered by the same circumstances, most whites who reach the more "difficult" level of having a black friend also have a black acquaintance. Thus, about 28 percent of whites with one or more black acquaintances also have at least one black friend, but fully 66 percent of whites with one or more black friends have at least one black acquaintance. In view of the heavy overlap between the two levels of interracial contact, we examine their effects on attitudes concurrently. This has the additional benefit of allowing us to examine their separate and joint effects. This is important because, unless whites have both friends *and* acquaintances who are black, their interracial contact is probably restricted to just one (token) black. About two-thirds of whites with interracial friendships have only one black friend, and 94 percent of those whites who do not describe their acquaintances as "all white" say they are "mostly white," which probably indicates no more than one black acquaintance. Are whites' racial attitudes more likely to be influenced when their personal contact with blacks exceeds a minimal, token level?

Table 1 reports the percentage of whites with conservative racial attitudes, according to whether they have at least one black friend and/or acquaintance. Responses are dichotomized into those with and with-

Table 1. Percentage of Whites with Conservative Racial Attitudes, by Whether They Have Black Friends and/or Black Acquaintances

	No Black Acquaintance		*Black Acquaintance(s)*	
	No Black Friend	*Black Friend(s)*	*No Black Friend*	*Black Friend(s)*
More whites dependable	53.8	51.0	46.6†	23.3**††
More whites intelligent	59.8	51.0	50.7††	38.6
More blacks lazy	38.7	28.0	32.9	26.8
Warmer toward whites	49.8	33.3*	35.8††	18.1**†
Closer toward whites	65.5	42.0**	51.5††	30.1**
Prefer whites in workplace	58.7	41.5**	38.9††	22.1**††
Prefer whites in neighborhood	84.9	76.9	69.0††	48.9**††
Government should do *same or less* to ensure:				
Integrated schools	84.2	75.5	83.7	77.2
Equal housing opportunity	73.0	65.9	65.7†	64.6
Equal job opportunity	79.5	72.5	72.1†	64.4
Base *N* (ranges)	925–1162	44–53	207–240	79–95

* Net difference between whites with and without *black friends* statistically significant ($p \leq .05$; chi-square test)

** Net difference between whites with and without *black friends* statistically significant ($p \leq .01$; chi-square test)

† Net difference between whites with and without *black acquaintances* statistically significant ($p \leq .05$; chi-square test)

†† Net difference between whites with and without *black acquaintances* statistically significant ($p \leq .01$; chi-square test)

out any black friends or acquaintances, respectively. Responses to the attitude items are also dichotomized, focusing on the conservative end of the distribution.[6]

Results indicate, first, that the effects of having a black acquaintance are about the same as the effects of a black friend. A comparison of the second column (black friend[s] but no black acquaintance) and third column (black acquaintance[s] but no black friend) reveals few differences, with no systematic pattern to the few small differences that do exist.[7] This suggests that it is not the closer intimacy of good friendship that is critical, but simply the experience with individual blacks that can come from relatively superficial relationships. Since whites are more likely to acquire black acquaintances than friends, the comparability of the effects of these two forms of personal contact on racial attitudes is encouraging.

---

[6] Because of the infrequency of personal interracial contact, there are too few whites in critical cells for us to conduct multivariate analyses that would control for the potential effects of demographic variables. After examining the effects of several demographic variables on (a) our measures of interracial contact and (b) our racial attitude measures, we determined that there were only two such variables that could potentially be confounding factors, region of residence (South/non-South) and age. Both demographic variables are weakly related to one or more of our contact measures as well as to racial attitudes. Small *N*'s preclude the examination of effects for southerners (i.e., whites residing in the former Confederate states), but all of our tables on the effects of racial contact (Tables 1, 2, and 3) have been recalculated for nonsouthern whites only, without any noteworthy alteration in the pattern or strength of the results. Results were also unaffected when we recalculated Tables 1, 2, and 3 for whites dichotomized into those aged less than 35 and 35 or older. Some may consider education another potentially confounding factor, but results reported by Jackman and Muha (1984) indicate that education is unrelated to the attitude measures used in the present article.

The limited effects of either form of contact, however, are less encouraging. While the size of the percentage differences fluctuates, the independent effects of either friend(s) or acquaintance(s) in the absence of the other are small. In general, whites must have both friend(s) *and* acquaintance(s) who are black before there is any appreciable impact on their attitudes toward blacks.[8] It seems more important to escape tokenism by establishing multiple contacts with blacks than to attain a high degree of personal intimacy with one's black contacts.

The association between multiple personal contacts with blacks and racial attitudes ranges from considerable to negligible, depending on the attitudinal element in question. The two aspects of whites' racial attitudes that seem to succumb most to the pressures of personal contact with blacks are affective preferences for whites over blacks and personal predispositions to avoid contact with blacks. The tendency to prefer whites over blacks in basic feelings of warmth and closeness and in personal social predispositions in the workplace and neighborhood does decline quite markedly, especially among whites who have both friends and acquaintances who are black.[9] The difference between the

[7] Because of the larger *N* with black acquaintances than with black friends, percentages in the third column are more likely to be statistically significant.

[8] We obtained the same pattern of results as that displayed in Table 1 with an alternative classification of black friends and acquaintances, as follows: (1) No Black Friends or Acquaintances, (2) One Black Friend *or* Mostly White Acquaintances, (3) More than One Black Friend *or* Acquaintances Half Black or More *or* Black Friend(s) and Acquaintance(s).

[9] On the social predisposition items, we consider the responses "All white" and "Mostly white" both as expressions of a preference for whites, as opposed to responses of half black or more or the expression of no preference. (Of course, since blacks constitute only about 12 percent of the population, it would be impossible to make all neighborhoods or places of works half black or more; however, the respondent is being asked for an expression of *personal* social predispositions toward blacks, not for a generalized policy feasibility assessment). While the response "All white" certainly represents a more blatant expression of racial preference than does "Mostly white," the distinction between these two responses may have more to do with ideological style than with the substantive issue of maintaining or abandoning white dominance (Jackman and Muha, 1984: 764). If we look at the "All white" and the "Mostly white" responses

latter group of whites and whites with no black friends or acquaintances ranges between 32 and 37 percent on these four items (second and third panels of Table 1).[10] Negative beliefs about blacks show less consistent declines: the difference between whites with no blacks in their social circles and whites with both black friends and black acquaintances is about 30 percent for the trait Dependable, 21 percent for the trait Intelligent, and 12 percent for the trait Lazy (first panel of Table 1). The perceived trait distinction that shows the smallest change (Lazy) is the one that has the smallest endorsement among whites initially. Nevertheless, the independent effects of either black friends or black acquaintances in the absence of the other type of contact are smaller for all three trait distinctions than for affective distinctions or personal social predispositions.

Finally, black friends and acquaintances have almost no effect on whites' policy orientations toward blacks. As one reads across the three rows in the bottom panel of Table 1, there is generally negligible change in the prevailing conservatism of whites' racial policy views.[11] This means that between two-thirds and three-quarters of whites who have blacks among their good friends and acquaintances persist in opposing any increase in governmental attempts to promote racial equality.

THE EFFECTS OF PERSONAL CONTACT VERSUS SHEER PROXIMITY

The development of personal relationships between whites and blacks is heavily determined by their physical proximity. Thus, our assessment of the effects of such personal relationships on racial at-

---

separately, they each show similar declines from one side of the table to the other: for Workplace, the "All white" response ranges in frequency from about 18 to 0 percent, the "Mostly white" response from about 40 to 22 percent; for Neighborhood, the "All white" response ranges from about 43 to 20 percent, the "Mostly white" response from about 41 to 28 percent.

[10] For the pairs of items reflecting both affective and social dispositions, one item is more "difficult" than the other. Whites are more reticent to express affective preference for their own racial group on the warm-cold dimension (which implies more overt hostility) than on the closeness dimension (the general tendency of whites to avoid expressing flagrantly hostile racial attitudes is discussed in Jackman and Muha, 1984). In their social dispositions, whites are less likely to express a preference for their own racial group in the workplace than in the neighborhood: the neighborhood is apparently more sensitively regarded as a bastion to be protected from black intrusion. Both of these distinctions are maintained across different levels of the Black Friends and Acquaintances variables.

[11] Many of those whom we term "conservative" are actually reactionary in their racial policy views: that is, they think the government should be doing *less* than it currently is to promote racial equality. On the issue of integrated schools, the view that the government should be doing *less* is expressed by 60 percent of whites with no black friends or acquaintances and 49 percent of whites with black friend(s) *and* acquaintance(s). On the other two issues, reactionary views are expressed by about 34 percent of the former group and 19 percent of the latter group of whites.

titudes must take stock of any possible effects of proximity itself. The contact theory emphasizes that the development of personal ties is critical. If so, controlling for sheer proximity to blacks should not alter the effects we observed in Table 1.

Hamilton and Bishop (1976) found a small positive change in the "symbolic racism" scores of whites who had acquired a black family in their neighborhoods in the previous year, whether or not they had ever interacted with that family. Wilner et al. (1952:67–68) found a more complex association whereby the degree of proximity to blacks and the amount of personal interaction with blacks were mutually reinforcing in bringing about a positive change in white housewives' "esteem" for blacks. The inconsistent results between these two studies may reflect differences in their attitude measures. Hamilton and Bishop's "symbolic racism" scale is more policy-oriented and also includes opinions about "symbolically laden" topics such as the safety of neighborhood streets, while Wilner et al.'s measure is of racial affect. Our data in Table 1 indicate that affect is more sensitive to the effects of personal contact. The difference in results may also stem from the two measures of personal interaction: Hamilton and Bishop's measure largely reflects relatively infrequent and possibly fleeting personal contact with a black neighbor, while Wilner et al.'s measure reflects repeated, sustained personal interaction. Our data permit us to resolve those inconsistencies by examining our array of racial attitude measures for the effects of residential proximity to blacks and sustained personal association with blacks.

Table 2 presents the percentage of whites with conservative racial attitudes, according to their degree of personal contact with blacks and their residential proximity to blacks. The results from Table 1 allow us to treat black friends and acquaintances as equivalent, and to combine them in the following trichotomy: (1) whites with No Black Friends or Acquaintances, (2) whites with Black Friend(s) *or* Black Acquaintance(s), and (3) whites with Black Friend(s) *and* Black Acquaintance(s). Table 2 is also divided into three vertical sections according to whites' residential proximity to blacks: (1) whites who have Never lived in a neighborhood with blacks, (2) whites whose Previous or Current neighborhood includes some blacks, and (3) whites whose Previous and Current neighborhoods include some blacks. This trichotomy focuses on how sustained whites' proximity to blacks has been over time.[12]

---

[12] Clearly, the neighborhood is not the only place where whites may come into contact with blacks, but it does constitute one major potential meeting place. Of the two arenas for which we have data on proximity to blacks, we have chosen to focus on the neighborhood rather than the workplace, for two reasons. First, data on workplace proximity are

Table 2. Percentage of Whites with Conservative Racial Attitudes, by Whether They Have Black Friends and Acquaintances and by Previous and Current Proximity to Blacks in the Neighborhood

	Proximity to Blacks in Neighborhood								
	Never			*Previous or Current*			*Previous and Current*		
	Black Friends/Acquaint.			*Black Friends/Acquaint.*			*Black Friends/Acquaint.*		
	None	*Friend* or *Acq.*	*Friend* and *Acq.*	*None*	*Friend* or *Acq.*	*Friend* and *Acq.*	*None*	*Friend* or *Acq.*	*Friend* and *Acq.*
More whites dependable	56.2	53.6	(25.0)	51.4	48.4	29.3*	53.4	40.2	16.2**
More whites intelligent	62.5	57.4	(50.0)	57.6	53.2	57.5	58.4	41.0	14.3**††
More blacks lazy	38.1	40.0	(37.5)	40.2	30.7	34.1	34.9	28.9	15.2
Warmer to whites	53.4	46.4	(22.2)	44.5	35.5	25.0**	54.4	27.7	9.8**
Closer to whites	67.4	58.2	(40.0)	64.1	50.4	29.5**	64.0	43.6	28.2**
Prefer whites in workplace	61.1	52.6	(20.0)*	58.3	41.6	34.1**	51.9	27.4††	9.8**†
Prefer whites in neighborhood	87.5	91.2	(55.6)**	84.6	72.3	65.9**	76.5††	54.3††	29.3**††
Government should do *same or less* to ensure:									
Integrated schools	83.3	86.8	(90.0)	84.3	86.3	83.7	85.6	73.7†	66.7**
Equal housing opportunity	73.1	78.3	(85.7)	73.3	66.7	71.8	70.7	56.5†	51.5*
Equal job opportunity	78.8	73.6	(88.9)	80.4	75.6	64.3*	77.8	65.6	59.0*
Base *N* (ranges)	394–506	46–56	7–10	390–482	117–138	39–44	133–162	83–95	33–41

* Net effect of *friends/acquaintances* statistically significant ($p \leq .05$; chi-square test)
** Net effect of *friends/acquaintances* statistically significant ($p \leq .01$; chi-square test)
† Net effect of *neighborhood proximity* statistically significant ($p \leq .05$; chi-square test)
†† Net effect of *neighborhood proximity* statistically significant ($p \leq .01$; chi-square test)

The effects of proximity and personal contact are mutually dependent and mutually reinforcing: each must be present for the other to have an effect, and the impact of each tends to increase as the level of the other increases. These results confirm and generalize for a broad array of racial attitudes the results reported by Wilner et al. (1952) for

---

available only for respondents who are currently employed, a restriction that severely reduces the *N* (unless we care to make the heroic assumption that people who are not employed have less proximity to blacks). The neighborhood is a universal experiential locale. Second, since whites are more resistant to the idea of having blacks in the neighborhood than in the workplace, the neighborhood represents a more significant arena of proximity. When we reestimate Table 2, substituting Workplace proximity for Neighborhood proximity, the *N*'s in three of the cells become too small, making it difficult to interpret the results with confidence. However, those results suggest that proximity to blacks in the Workplace has less import for whites' racial attitudes than does proximity to blacks in the Neighborhood.

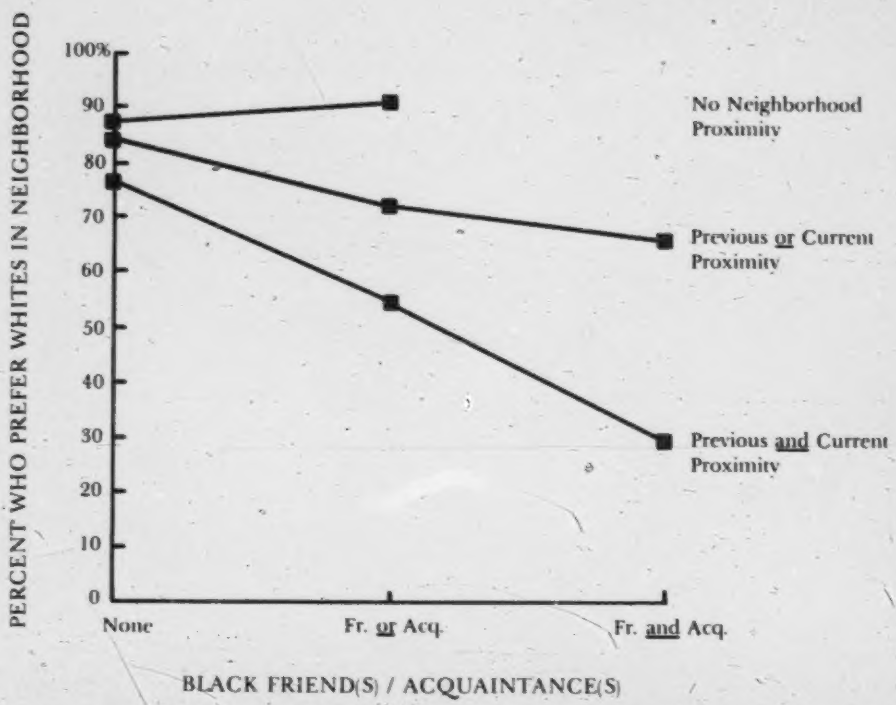


Figure 1. Effects of Black Friends and Acquaintances on Whites' Expressed Preference for Whites in the Neighborhood, by Different Levels of Proximity to Blacks in the Neighborhood

racial "esteem." Sheer proximity to blacks appears to be of little value, unless it is accompanied by personal contact, but proximity does have a direct effect of its own on racial attitudes when personal contact accompanies it, and the more personal contact there is, the greater the effect of proximity. Conversely, it appears that personal contact needs to be backed up by physical proximity to blacks if it is to influence whites' racial attitudes, and the more sustained the proximity, the greater the impact that personal contact has. These conclusions must be tempered somewhat in view of the virtually empty cell for whites who have high levels of personal contact with blacks but no history of proximity to blacks in their neighborhoods, but confidence is encouraged by the consistency of the pattern across all other cells in Table 2. These results reinforce our inference from Table 1 that diversity of contacts with blacks is more critical than the intimacy of interracial contacts. The pattern of results is illustrated in Figure 1.

If one compares the net effects of personal contact in Table 2 with those reported in Table 1, the relative susceptibility of various racial attitudes remains unchanged. As before, affective and personal predispositions are generally the most strongly related to personal contact. However, in the highest proximity condition, contact effects on two of the three trait distinctions are just as large. And, unlike the effects in Table 1, there is consistent, significant movement in policy orientations as personal contact increases. The upper-bound net effects of residential proximity are harder to assess, because of the empty cell in the low-proximity/high-contact condition. However, it is personal predispositions for contact with blacks that seem to be the most clearly susceptible to the effects of residential proximity, especially among whites with only token personal contact with blacks. The net impact of residential proximity to blacks generally appears to be comparable to the net impact of personal contact with blacks. Personal contact has slightly more consistent effects on affective and social predispositions, while residential proximity has a somewhat stronger effect on policy orientations.

The *joint* effects of residential proximity to blacks and personal contact with blacks represent a considerable advance over the effects of personal contact reported in Table 1. Cognitive distinctions (with the exception of the trait Lazy, which is less popularly endorsed) and affective distinctions show a total change from one side of the table to the other of about 40 percentage points or more, while social predispositions indicate an even larger change of about 50 to 60 percentage points. The total change in policy orientations lags considerably behind, with the percentage difference generally about 20 points from one side of the table to the other. Overall, the inclusion of residential proximity to blacks pushes attitudes further away from racial conservatism by an additional 10–26 percentage points over the total movement in attitudes wrought by black friends and acquaintances in Table 1.[13]

Yet even the joint effects of multiple interracial experiences are limited. The two attitude items that are influenced most by the joint effects of residential proximity and personal contact are those that are substantively closest to the specifics of the racial contact itself: personal racial predispositions in the neighborhood (largest effect) and in the workplace (second largest effect). Changes in orientations toward governmental policies to promote racial equality, while more appreciable than in Table 1, still remain modest. The net result is that among the select group of whites with sustained residential proximity to blacks and with blacks among their friends and acquaintances, the proportion with negative racial beliefs, feelings, and social predispositions is reduced to a small minority—and yet resistance to policies affirming racial equality is still found among half to two-thirds of this group.

---

[13] The only two exceptions to this statement are the items "Closer to whites" and "Equal job opportunities," both of which evidence about the same degree of movement in Tables 1 and 2, although the sources of that movement are apportioned differently.

THE RELATIVE SOCIOECONOMIC STATUS OF BLACK FRIENDS

A common implicit assumption has been that the ties of friendship confer an instrinsic equality between the participants, at least within the bounds of the relationship. However, while the relationship itself does not manufacture an inequality between the participants (as in master–servant relationship), it is also possible that friendship may do nothing to erase an inequality that exists on the basis of socially important characteristics of the participants, such as race or socioeconomic status. Such a possibility is suggested by our results thus far, which indicate that interracial friendship does not have especially strong effects on racial attitudes and that support for policy efforts to promote racial equality are the most resistant to change. The pertinent issue may not be whether a personal relationship of equality is generalized to the group as a whole, but rather whether the black friend was ever an equal in the first place.

There is no way to ascertain definitively that any of the interracial friendships involve relationships between equals. However, other status characteristics of the individual friend may offset his racial status. Of these, the strongest and most visible candidate is the friend's socioeconomic status. In the classic public housing studies of Deutsch and Collins (1951) and Wilner et al. (1952), the socioeconomic status of the participants was held constant, because of the income requirements for eligibility for public housing. Thus, the residents of both races shared not only the same rights and privileges as tenants within the project, but also the same (low) socioeconomic standing, and Wilner et al. (1952: 69) identified this as an important factor in the creation of conditions conducive to attitude change.

To pursue this issue, we can subdivide the select group of whites who have black friends by the relative socioeconomic status of those friends. The frame of reference for assessing this is the mean socioeconomic status of the respondent's white friends. We assume that people are cognizant of status differentials among their friends, and that when they evaulate the traits of any individual friend the most salient reference group is the remainder of their friendship circle.[14]

Respondents were asked to list the occupation of each of their friends, and we used that information to derive a Duncan SEI score for each friend.[15] This allows us to compare the SEI score of the respondent's black friend (or the mean black SEI score if there is more than one black friend) with the mean SEI score of that respondent's white friends. On this basis, whites with black friends are divided into three groups: those whose black friends have *lower* socioeconomic status than their white friends, those whose black friends and white friends have *similar* SEI scores (that is, within 8 SEI points), and those whose black friends have *higher* socioeconomic status than their white friends.[16]

---

[14] The only other potential reference would be the respondent's own socioeconomic status, but use of this is impractical because the resulting measure would suffer from floor and ceiling effects, depending on whether the respondent's own status was low or high. Such a measure would thus artifactually reflect the status of the respondent as well as of the friend. Since the status of friends is strongly associated with respondent's own status (Jackman and Jackman, 1983), use of the absolute status of respondent's black friends would also be confounded with respondent's own status.

If status differentials do permeate the bounds of good friendship, we should expect the impact of interracial friendship on whites' racial attitudes to be heavily dependent on the relative socioeconomic status of their black friends. Such a pattern would given an important clue to the restricted effects of personal interracial contact on whites' racial attitudes. It would suggest that social inequality between groups has a potency that is not erased by personal friendship.

Table 3 displays the percentage with conservative racial attitudes, first of whites who do not have a black friend, then of whites who do have black friends, with the latter group subdivided according to whether their black friends' socioeconomic status is lower than, the same as, or higher than that of their white friends. In view of the small *N*'s in the last three categories, we must interpret the results in Table 3 with caution. However, the patterns are remarkably consistent and pronounced.[17]

The relative status of black friends is a consistent influence on all aspects of whites' racial attitudes, even their policy orientations. Its

---

[15] Data provided by respondents about their friends' occupations were sufficiently detailed to allow for the Survey Research Center two-digit occupation code, which is an abbreviated version of the three-digit 1970 Bureau of the Census occupation code. For each friend, we calculated a Duncan SEI score by taking the weighted mean SEI of the three-digit census occupations that were included in each two-digit SRC occupation category (see Jackman and Jackman, 1983, for more details). Individual friends were excluded if their occupational status was not ascertained.

[16] Not surprisingly, the mean SEI differential in the Black Friend Lower Status category is greater (mean difference = 26 SEI points) than in the Black Friend Higher Status category (mean difference = 18 SEI points) In view of the small *N* and the arbitrary cutoffs between categories (± 8 SEI points), we conducted two checks. First, we experimented with two slightly different categorizations, one based on a different score of ± 7 SEI points, and one based on ratios of less than .85 and more than 1.15. The *N*'s changed slightly with each categorization, but the results were replicated. Second, we hand-checked the questionnaires of the 96 white respondents with valid occupational data for black and white friends. This check indicated that the three categories for the relative SEI of black and white friends reflect plausible, coherent distinctions.

[17] As with the controls for region and age (see footnote 6 above), the results in Table 3 are not undermined when we control for whether respondents have both sustained proximity to blacks in the neighborhood *and* black acquaintance(s) as well as friend(s).

Table 3. Percentage of Whites with Conservative Racial Attitudes, by Whether They Have Black Friends and by the Relative Socioeconomic Status of Their Black and White Friends

	No Black Friend	*Black Friend Lower SEI*	*Black Friend Same SEI*	*Black Friend Higher SEI*
More whites dependable	52.5	43.8	36.1	19.0**
More whites intelligent	58.1	54.8	44.4	28.6*
More blacks lazy	37.7	40.0	27.8	14.3
Warmer toward whites	47.3	32.4	33.3	4.8**†
Closer toward whites	63.0	47.1	44.7	10.0**††
Prefer whites in workplace	55.3	34.3	28.2	27.3**
Prefer whites in neighborhood	82.2	73.5	61.5	36.4**†
Government should do *same or less* to ensure:				
Integrated schools	84.2	91.2	77.8	61.9*†
Equal housing opportunity	71.7	86.7	67.7	52.6†
Equal job opportunity	78.3	85.3	63.2	47.6**††
Base *N* (ranges)	1136–1406	30–35	31–39	19–22

* Significantly different from *No Black Friend* category ($p \leq .05$; chi-square test)

** Significantly different from *No Black Friend* category ($p \leq .01$; chi-square test)

† Significantly different from *Black Friend Lower SEI* category ($p \leq .05$; chi-square test)

†† Significantly different from *Black Friend Lower SEI* category ($p \leq .01$; chi-square test)

importance can be seen in two ways. First, having a black friend of relatively low socioeconomic status generally does little to temper whites' attitudes toward blacks. If one compares the percentages in the first two columns, the differences between whites with no black friends and whites with black friends of relatively low socioeconomic status tend to be small and irregular. Second, if one compares the percentages across the three categories of black friends' socioeconomic status, one finds an almost unbroken pattern of substantial positive change in whites' racial attitudes as the relative socioeconomic status of their black friends increases. The differences in racial attitudes between whites with lower- and higher-status black friends are generally as large as or larger than the differences observed in Table 1 between whites with no black friends or acquaintances and whites with both black friends and acquaintances. The only discrepancy is in the tendency to prefer whites in the workplace, which is more influenced by the friendship factor itself and little affected by the status of the friend. On all the other items, the percentage difference between whites with lower- and higher-status black friends ranges between 25 and 38 percentage points.

In the tendency to make negative trait distinctions, the difference is about 25 percentage points for each trait, which is not quite large enough to attain statistical significance at the .05 level with the small cell *N*'s. As a result of the consistent positive shift in trait attributions, however, only about one-seventh to one-quarter of whites with higher-status black friends make negative trait distinctions between the races. In affective differentiation between the races, in personal preferences for whites in the neighborhood, and in racial policy orientations, the differences between the second and fourth columns of Table 3 all range between 28 and 38 percentage points, and these differences are large enough to attain statistical significance at the .05 level or better. As a result, negative affect toward blacks is almost obliterated among whites with higher-status black friends, while about one-third of this group maintain a preference for whites in the neighborhood, and about one-half maintain a conservative racial policy stance. The last effect is especially noteworthy, in view of the limited effects on racial policy views observed in our earlier analyses. While almost all whites with a lower-status black friend preserve a withholding policy perspective toward blacks, there is a clear movement away from that position as their black friends have higher status. The remaining high proportion of whites with conservative racial policy views indicates, however, that it would take more than one (token) high-status black friend to achieve a radical shift in the overwhelming conservatism of whites' racial policy views.

Discussion

Our results suggest several points of divergence from the contact theory. First, while the importance of developing a variety of contacts with blacks is congruent with the contact theory, the same cannot be said of the apparent unimportance of the intimacy of those contacts. The contact theory emphasizes that contacts with duration and intimacy (such as friendships) are more motivationally compelling. We find no evidence of that. Second, the contact theory portrays prejudice and discrimination as emotionally based outgrowths of the ignorance that accompanies the physical separation of blacks and whites. From this perspective, prejudice may be seen as a tragic misunderstanding that can be set right by introducing compelling cross-group contacts. Our results do not support such an optimistic view. Instead, they suggest that whites' affective and social dispositions toward blacks change with greater ease than their beliefs about blacks, or, more dramatically, their racial policy views. Finally, the significance of the socioeconomic status of even good friends who are black is broadly consistent with the stipulation of the contact theory that interracial contacts should take place between status equals to be effective. However, when placed in the context of our other results, and especially the resilient conservatism of whites' policy dispositions toward blacks, the meaning of the socioeconomic status factor changes. Rather than seeing this as something that can render blacks and whites equals in a particular situation or relationship, we believe its significance is that it can offset the status differential that is embedded on the basis of race. This interpretation is reinforced by the fact that racial attitudes are more positive when black friends have *higher* socioeconomic status than when they have equal status. This suggests that the critical dynamic is not that blacks and whites are rendered equals but rather that another dimension of inequality may partially balance off the nested racial inequality. That such an extrinsic factor invades even the bounds of intimate friendship reveals the pervasive force of societally defined inequalities between groups in determining the predispositions of individuals.

The policy implications of these results are less than encouraging. The lack of necessity for highly intimate contacts across racial lines is a plus, but the importance of experiencing a variety of interracial contacts is a serious drawback, since most whites who do have contact with blacks experience only token contact. Even more discouraging is the apparently critical significance of the relative socioeconomic status of black contacts. It appears that unless an increase in interracial contact is accompanied by wide-scale change in the relative socioeconomic position of blacks, it is unlikely to have a salutary effect on whites' racial policy views. Paradoxically, however, the government is unlikely to push hard for affirmative change in the relative socioeconomic standing of blacks when its primarily white constituency is so overwhelmingly opposed to such change. Stated differently, a relationship between whites and blacks that was not marked by discrimination and socioeconomic inequality would be a very different relationship from the one with which we are confronted in the United States today. The important question is not whether amicable contact can exist between groups of equal status, but whether such contact can help to foster equality between groups who have unequal status.

We propose that a more political conception of intergroup attitudes is in order. Instead of viewing such attitudes primarily as individual expressions of irrationality, parochialism, and animosity, we suggest they are a constituent part of a group's ideological defense of its interests. As the dominant group develops an ideological defense of the inequality from which it benefits, the natural inclination is to mitigate incipient challenge from subordinates through the use of amicable persuasion, rather than to heighten the conflict with a mindless reliance on hostile negativism (Jackman and Muha, 1984). From this perspective, the empirical relationship between interracial contact and whites' racial attitudes bears on a different issue than that posed by the contact theory. The issue is not whether whites generalize to blacks as a whole from their personal experience of friendship with an individual black. Instead, the issue is how a relationship of intimacy with individual subordinates modifies the manner in which dominant-group members defend their privilege.

The idea that intergroup attitudes are energized by a collective sense of group position, rather than being simple reflections of personal, experiential factors, was suggested long ago by Blumer: "The sense of group position is the very heart of the relation of the dominant to the subordinate group. . . . To seek . . . to understand it [race prejudice] or to handle it in the arena of individual feeling and of individual experience seems to me to be clearly misdirected" (Blumer, 1958: 4, 6–7). Such a view diverges sharply from the more prevalent view of prejudice expressed in the contact theory, and, indeed, Blumer's argument went largely unattended in empirical research on intergroup attitudes. In recent years, however, some analysts have begun to pursue empirical questions that are informed by the idea of intergroup attitudes as an expression of group-level interests (e.g., Wellman, 1977; Smith, 1981; Bobo, 1983, 1984; Jackman and Senter, 1983; Jackman and Muha, 1984).

If prejudice derives not from feelings of personal animosity but from an implicit sense of group position, then dominant groups will seek to defend their privilege no matter what brand of affect they feel toward subordinates. This, in turn, implies that there is no logical connection between friendship and equality. Hacker (1951) made precisely this point about gender relations in her classic exposition on that subject. She argued that personal friendship between men and women does not portend sexual equality, and that, indeed, marriage (the ultimate point of acceptance on Bogardus' classic social distance scale) is an intimate relationship of inequality. We suggest that personal geniality may have been overrated in the analysis of racial attitudes as well. In the segregation-charged race relations of recent American history, it is easy to forget that racial privilege can be maintained as adeptly with personal relationships as with impersonal ones. The corollary is that discriminatory orientations need not rest on feelings of personal animosity. Indeed, if circumstances are conducive, discrimination can rest as comfortably on a friendly disposition as on a hostile one.

When whites experience a high degree of personal contact with blacks, feelings of personal animosity and social distance crumble. This, however, does not alter the relationship of inequality that exists between the two racial groups and from which whites benefit. Thus, what is wrought by interracial contact is a significant modification in the *configuration* of whites' racial attitudes—indeed, they begin to resemble the configuration of men's gender attitudes. Men, who almost uniformly experience high levels of personal contact with women, tend to express positive affective dispositions toward women, while they withhold support for the promotion of women's equality. Data from the same national survey from which our racial attitude data were drawn indicate that very few men express an affective preference for men over women (in terms of either warmth or closeness), but almost two-thirds think that the government should not increase its efforts to promote equality between the sexes in legal rights or job opportunities. This is paralleled by the profile of racial attitudes among whites who have high levels of contact with blacks: that contact renders a substantial positive change in whites' personal affective and social predispositions without sweeping away discriminatory policy orientations. The latter are driven by an enduring force that is unaltered by a change in personal contact with blacks—the material and cultural interests of white racial privilege. As women have long understood implicitly, intergroup friendship increases the bonds of affection with subordinates, but it does not undercut the discrimination that defines the unequal relationship between the two groups.

Appendix A

MEASURES OF RACIAL ATTITUDES

Cognitive Differentiation Between Blacks and Whites. In a self-administered booklet, respondents were asked three pairs of questions about the trait attributes of (a) blacks and (b) whites:

> How many [blacks/whites] would you say are [dependable/intelligent/lazy]?

Respondents answered by writing a number from a scale that ranged from 0 (None) through 4 (About Half) to 8 (All). We took the difference between each respondent's answer for blacks and for whites to identify respondents who think there are *more dependable whites* than blacks, *more intelligent whites* than blacks, and *more lazy blacks* than whites.

Affective Differentiation Between Blacks and Whites. These measures are also constructed from pairs of items asked in a self-administered booklet:

> In general, how warm or cold do you feel toward [blacks/whites]?
> In general, how close do you feel to [blacks/whites]?

Respondents answered by writing a number from a scale that ranged from 1 (Very Cold/Not at all Close) through 5 (Neither One Feeling Nor the Other) to 9 (Very Warm/Very Close). We took the difference between each respondent's answer for blacks and for whites to identify respondents who feel *warmer toward whites* than to blacks, and those who feel *closer to whites* than to blacks.

Social Predispositions Toward Blacks. These measures come from two items:

> Would you personally prefer to work in a job where your co-workers are all white, mostly white, about half white and half black, mostly black, or all black?

> Would you personally prefer to live in a neighborhood that's all white, mostly white, about half white and half black, mostly black, or all black?

Those who Prefer Whites in the Workplace or Prefer Whites in the Neighborhood are those who responded "all white" or "mostly white" to the appropriate question. Volunteered "No Preference" and "Don't Know" responses are coded as nondiscriminatory.

Racial Policy Orientations. These measures come from three pairs of questions:

> How much do you think the federal government *is* doing to [policy statement]? A lot, Quite a Bit, A Little, or Nothing?
>
> How much do you think the federal government *should be* doing about this? A Lot, Quite a Bit, A Little, or Nothing?

The three Policy Statements are:

> Make sure that blacks and whites go to the same schools.
> Make sure that blacks can buy any house on the market that they can afford.
> Make sure blacks have the same job opportunities as whites.

We took the difference between each respondent's answers to the two questions in each pair to identify those who think the government *should be doing less* than it is now, the *same* as it is doing now, or *more* than it is doing now to promote *integrated schools, equal housing opportunities,* and *equal job opportunities*. Respondents with "Don't know" or missing data on either item in a pair were excluded.

NOTE: Use of the policy "should" items alone, instead of the policy difference scores, does not alter the pattern of results reported in this article.

Appendix B

SUPPLEMENTARY EVIDENCE ON SELF-SELECTION VERSUS UNSOLICITED PROXIMITY TO BLACKS

1. *Proximity and Friendship*. Table 4 displays, for white respondents, the level of informal interaction with blacks in the neighborhood, the racial composition of acquaintances, and the number of black friends, as a function of the racial composition of the current neighborhood of residence and the presence or absence of blacks in any previous neighborhood of residence. Comparable data are available for the workplace, and they yield similar results to those reported for the neighborhood, the only difference being the generally higher levels of interracial interaction in the workplace (which is consistent with our view that the workplace facilitates social interaction more than does the neighborhood).

The results in Table 4 suggest that once whites find themselves in an interracial situation, their patterns of social interaction gradually succumb to the intrinsic pressures of prolonged and frequent exposure to blacks, according to the same principles that determine sociometric patterns more generally. The strength of the relationship is underscored when one considers the partial nature of our measures of proximity: the actual physical proximity and accessibility of black families in the neighborhood would vary considerably, with many cases presenting only limited social opportunities.

The effects of both current and previous exposure to blacks on the likelihood of engaging in neighborly activities with blacks in the current neighborhood of residence suggest a cumulative process whereby whites with more interracial experience find it easier to establish contact when they encounter blacks subsequently. The combined effects of previous and current neighborhood proximity produce a large percentage difference between the two "ends" of the table. Among whites whose current neighborhood is "mostly white" and who have not previously lived in a neighborhood with any black residents, only 18 percent ever engage in neighborly activities with their black neighbors. In contrast, among whites with both some past neighborhood proximity and a current neighborhood that is half black or more, almost 60 percent interact with their black neighbors. The comparable figures for workplace proximity are 45 percent and 80 percent.

Table 4. Percentage of Whites Who Interact with Blacks, as a Function of Past Exposure to Blacks in the Neighborhood and Racial Composition of Current Neighborhood

	Previous Neighborhood No Blacks			*Previous Neighborhood One or More Blacks*		
Current Neighborhood	*All White*	*Mostly White*	*Half Black or More*	*All White*	*Mostly White*	*Half Black or More*
Neighborly activities w/blacks in neighborhood						
Never	—	82.0	67.9	—	72.4	40.5
Not too often	—	10.4	3.6	—	12.4	24.3
Sometimes	—	5.5	21.4	—	9.5	21.6
Very often	—	2.2	7.1	—	5.7	13.5
N	0	183	28	0	210	37
Race of acquaintances						
All white	89.7	73.0	71.4	80.6	61.3	32.4
Mostly white	10.0	24.3	25.0	18.5	34.9	59.5
Half black or more	0.3	2.7	3.6	0.9	3.8	8.1
N	602	226	28	438	261	37
Black friends?						
None	96.4	87.7	78.6	90.4	82.7	80.6
One	2.5	9.1	10.7	7.1	9.2	8.3
Two or more	1.0	3.2	10.7	2.5	8.1	11.1
N	590	220	28	436	260	36

The likelihood of interracial acquaintanceship is heavily determined by proximity to blacks. Among whites with no past or present proximity to blacks in the neighborhood, very few—only 10 percent—have any black acquaintances: this figure jumps to 68 percent among whites with both some previous neighborhood proximity to blacks and current residence in a neighborhood that is half black or more (second panel of Table 4). The comparable figures for workplace proximity are 7 percent and 44 percent. The influence of neighborhood proximity on interracial friendship seems less pronounced, although if one takes account of the small proportion of whites overall with black friends, the relationship is far from trivial. Ranging from 3.5 percent in the extreme lefthand column of Panel 3 to 19.4 percent in the extreme righthand column, this represents a ratio of almost 6 to 1 in favor of having a black friend at higher versus lower levels of neighborhood proximity. In the work context, the comparable percentages are 2 and 26 (or a ratio of more than 12 to 1 favoring interracial friendships at high proximity levels).

2. *Interracial Contact and Racial Attitudes*. Several aspects of the relationship between interracial contact and whites' racial attitudes contradict the logical expectations of the self-selection thesis, and thus reinforce our assumption that interracial friendships develop primarily from factors that are independent of initial racial attitude. We briefly identify the pertinent results from Tables 1, 2, and 3.

i. The overall relationship between our various measures of interracial contact and racial attitudes is neither strong nor uniform enough to provide support for the self-selection thesis. Tables 1 and 2 indicate that a significant portion of whites with high levels of interracial contact have negative racial attitudes. This contradicts the idea that whites with positive attitudes seek personal contact with blacks or that whites with negative attitudes avoid such contact.

ii. Table 1 indicates that the relationship between contact and attitudes is the same, whether the contacts are casual acquaintances or more intimate friends. If self-selection bias were operating, whites with black friends would have more positive racial attitudes than those whose relationships with blacks did not extend beyond casual acquaintance.

iii. The mutually reinforcing effects of proximity and personal association reported in Table 2 undermine the idea that whites with negative racial attitudes either avoid neighborhoods with blacks in them or avoid interacting with blacks in those neighborhoods. If the first type of avoidance were taking place, negative racial attitudes would be absent or near absent among all white residents of interracial neighborhoods. If the second type of avoidance were taking place, the white residents of interracial neighborhoods who make *no* personal contact with blacks would have more *negative* racial attitudes than a cross-section of whites who have had no exposure to blacks (because those in the interracial neighborhood with positive attitudes would have self-selected into personal relationships with blacks). The data in Table 2 offer no evidence for either of these expectations. Instead, whites who experience either proximity or personal contact without the reinforcement of the other have essentially the same profile of racial attitudes as whites with no exposure to blacks.

iv. The relationships reported in Table 3 between the socioeconomic status of black friends and whites' racial attitudes undermine the plausibility of the self-selection thesis. It is implausible to posit that whites' racial attitudes determine the relative socioeconomic status of the blacks with whom they become friendly. The implausibility of such a proposition is underscored by the fact that whites with lower-status black friends do not display systematically more negative attitudes than whites with no black friends (that is, there is no evidence that whites "sort" into high- and low-status friendships according to whether their racial attitudes are more positive or negative than average).

References

Allport, Gordon W.
1958 The Nature of Prejudice. New York: Anchor.
Amir, Yehuda
1969 "Contact hypothesis in ethnic relations." Psychological Bulletin 71(May):319–42.
Berscheid, Ellen, and Elaine Hatfield Walster
1969 Interpersonal Attraction. Reading, MA: Addison-Wesley.
Blumer, Herbert
1958 "Race prejudice as a sense of group position." The Pacific Sociological Review 1(Spring):3–7.
Bobo, Lawrence
1983 "Whites' opposition to busing: symbolic racism or realistic group conflict?" Journal of Personality and Social Psychology 45:1196–1210.
1984 Racial Hegemony: Group Conflict, Prejudice, and the Paradox of American Racial Attitudes. Unpublished doctoral dissertation, Department of Sociology, University of Michigan.
Brophy, I. N.
1945 "The luxury of anti-Negro prejudice." Public Opinion Quarterly 9:456–466.
Cook, Stuart W.
1963 "Desegregation: a psychological analysis," *in* W. W. Charters, Jr., and N. L. Gage (eds.), Readings in the Social Psychology of Education. Boston: Allyn and Bacon.
Deutsch, Morton, and Mary Evans Collins
1951 Interracial Housing: A Psychological Evaluation of a Social Experiment. Minneapolis: University of Minnesota Press.
1965 "The effect of public policy in housing projects upon interracial attitudes." Pp. 646–657 *in* Harold Proshansky and Bernard Seidenberg (eds.), Basic Studies in Social Psychology. New York: Holt, Rinehart, and Winston.
Farley, Reynolds, Toni Richards, and Clarence Wurdock
1980 "School desegregation and white flight: an investigation of competing models and their discrepant findings." Sociology of Education 53 (July):123–139.
Festinger, Leon
1951 "Architecture and group membership." The Journal of Social Issues 7:152–163.

Festinger, Leon, Stanley Schachter, and Kurt Back
1950 Social Pressures in Informal Groups: A Study of Human Factors in Housing. New York: Harper and Brothers.
Hacker, Helen Mayer
1951 "Women as a minority group." Social Forces 30:60–69.
Hallinan, Maureen T.
1982 "Classroom racial composition and children's friendships." Social Forces 61 (September):56–72
Hallinan, Maureen T., and Stevens S. Smith
1985 "The effects of classroom racial composition on students' interracial friendliness." Social Psychology Quarterly 48(March):3–16.
Hamilton, David L., and George D. Bishop
1976 "Attitudinal and behavioral effects of initial integration of white suburban neighborhoods." The Journal of Social Issues 32 (2):47–67.
Hodge, Robert W., and Donald J. Treiman
1968 "Class identification in the United States." American Journal of Sociology 73 (March):535–47.
Jackman, Mary R., and Robert W. Jackman
1983 Class Awareness in the United States. Berkeley and Los Angeles: University of California Press.
Jackman, Mary R., and Michael J. Muha
1984 "Education and intergroup attitudes: moral enlightenment, superficial democratic commitment, or ideological refinement?" American Sociological Review 49 (December):751–69.
Jackman, Mary R., and Mary S. Senter
1983 "Different, therefore unequal: beliefs about trait differences between groups of unequal status." Pp. 309–35 in Donald J. Treiman and Robert V. Robinson (eds.), Research in Social Stratification and Mobility, vol. 2. Greenwich, CT: JAI Press.
Marx, Karl
1964 Karl Marx: Selected Writings in Sociology and Social Philosophy, edited and translated by T. B. Bottomore. New York: McGraw-Hill.
Myrdal, Gunnar
1944 An American Dilemma. New York: Harper and Brothers.
Polsby, Nelson W.
1980 Community Power and Political Theory, second edition. New Haven, CT: Yale University Press.
Smith, A. Wade
1981 "Racial tolerance as a function of group position." American Sociological Review 46 (October):558–73.
Stouffer, Samuel A., and others.
1949 The American Soldier, volume 1. Princeton, NJ: Princeton University Press.
de Tocqueville, Alexis
1969 Democracy in America. New York: Anchor Books.
Vander Zanden, James W.
1984 Social Psychology. New York: Random House.
Wellman, David T.
1977 Portraits of White Racism. New York: Cambridge University Press.
Williams, Robin M., Jr.
1964 Strangers Next Door: Ethnic Relations in American Communities. Englewood Cliffs, NJ: Prentice-Hall.
Wilner, Daniel M., Rosabelle Price Walkley, and Stuart W. Cook
1952 "Residential proximity and intergroup relations in public housing projects." The Journal of Social Issues 8:45–69.
1955 Human Relations in Interracial Housing. Minneapolis: University of Minnesota Press.

Yuppies

JOHN L. HAMMOND

THE MEDIA have reported the emergence of a new political species: the Young Urban Professionals, children of the baby boom who now have college degrees and high-paying jobs. The huge birth cohorts of the baby boom were said to share some important formative experiences; childhoods of prosperity, high educational levels, and now, occupational attainments and salaries which match their ambitions. According to press accounts, they crowd the country's labor and real estate markets, health clubs, and boutiques. A leading newsweekly proclaimed 1984 their "year" (Newsweek, 1984), and declared that they were becoming the dominant political and cultural force in American society.

Their politics appears to be new, too—conservative, but not on all issues. Media accounts, though impressionistic, have suggested that yuppies reject high taxes, social spending, and government regulation, the New Deal economic agenda. But they are liberal on issues of personal freedom and lifestyle, especially sexual and gender freedom: they support equal rights for women; they reject restrictive sexual mores, and accordingly endorse freedom of sexual preference and

---

Abstract According to widespread press reports, the Young Urban Professionals—members of the baby boom generation with college educations and high-paying jobs—are becoming a dominant political and cultural force in American society. While accounts of their attitudes are impressionistic, most suggest that they adopt liberal positions on issues of personal freedom but reject the socioeconomic liberalism of the New Deal and the 1960s. Analysis of the General Social Surveys from 1982 to 1984, however, shows that while the presumptive yuppies are more liberal than the general population on issues of personal freedom, they are not particularly conservative on social welfare issues. The differences are explained by their youth and higher education rather than by any distinct yuppie attitude syndrome. Some possible causes of the mass media's infatuation with a perhaps mythical yuppie are discussed..

John L. Hammond is Professor of Sociology at Hunter College and the Graduate Center of the City University of New York. The author expresses appreciation for the helpful comments of Al Gollin, Bernie Lefkowitz, Pam Stone, Gaye Tuchman, and Mike Wood.

Public Opinion Quarterly Vol. 50:487–501 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-487/$2.50

abortion rights. Finally, the media have pictured yuppies as materialistic, but not out of economic insecurity; rather, they are economically privileged and they take advantage of it.

But there is precious little evidence for this characterization, despite its widespread acceptance. As I will show, data from the recent General Social Surveys demonstrate that yuppies are as liberal, or more so, as the rest of the population on both lifestyle and economic issues. I will then consider why the unsubstantiated view has been so widely reported.

Sixties Going on Eighties

Those who attribute a unique pattern of beliefs to the yuppies argue that a number of trends in American culture generally and in baby boomers' lives converged to create it. Cultural restrictions have loosened. Many accounts suggest that as confidence in national institutions has been declining for two decades, personal fulfillment has become a dominant theme of the culture. People today are much more likely to find satisfaction in self-expression than in fulfilling social roles or adhering to institutional requirements (Bellah et al., 1985; Turner, 1976; Veroff et al., 1981; Yankelovich, 1981). These preoccupations have spawned a plethora of self-help movements and new forms of psychotherapy.

This loosening of cultural restraints is manifest in feminism: the belief in equal rights for women, the breakdown of traditional gender roles, and the access of women to new, if not precisely equal, opportunities. Closely related is the revolution in sexual mores: the apparent decline in the double standard and the growing belief that sexual preference (for partners of the same or the opposite sex, for one or many partners, for the traditionally sanctioned or more adventurous forms of sexual contact) is a matter of individual choice.

These new norms burst forth in the 1960s as part of a movement which was antiauthoritarian and politically left. The term *yuppie* resonates, after all, because it evokes the Yippies who nominated a pig for president in 1968. The affluent Flower Children, encouraged to believe that they could accomplish whatever they wanted, demanded freedom from restrictive norms. But they also demanded equal opportunity and a decent standard of living for minority groups, and fuller democratic and communal participation in decision making; they rejected wealth as the standard of social status. Some suggested that a "postmaterialist" value system was emerging appropriate to a postindustrial society where material needs were largely met (Inglehart, 1977).

But media accounts suggest that yuppies are different: they are nonconforming and antiauthoritarian, but they are also very concerned with income, consumption, and possessions. As Jones puts it, it is "as if the Woodstock mentality had somehow merged with the Bloomingdale mentality" (1981: 317).[1] Moreover, as mentioned above, they are said to oppose the economic liberalism of the New Deal. While commentators do not say so directly, many imply that yuppies' materialism and their economic conservatism go together: they reject New Deal liberalism *because* they are materialistic and self-serving (Boaz, 1985; Newsweek, 1984: 28–31). In a decade, they seem to have passed from the Greening of America to Looking Out for Number One.

Several events in the national political culture and in their own lives might have brought about this change. Declining trust in political institutions, generally attributed to the Vietnam War and to Watergate, may have led people to oppose all forms of government intervention. While these events affected everyone, they presumably affected the young most, because young people were getting their first exposure to politics. Further, traditional liberalism has lost its appeal. Social change, activist government, and the welfare of the underprivileged were long felt to go together. For many, they no longer do (Abramson, 1983: 181–82, 233–34; Jennings and Niemi, 1981: 203–11, 226; Mannheim, 1969; Smith, 1981).

The present apparent economic conservatism of the baby boom generation may derive from disappointed economic expectations, just as the behavior and beliefs of flower children grew out of affluence. Stagflation subverted liberalism's promise of unending prosperity, and competition for jobs was intense because their cohort was so big. Easterlin argues that the size of a birth cohort affects its members' life chances, and consequently their attitudes and behavior. For the members of a large birth cohort, competition on entry to the labor market is so intense that they become pessimistic, marry late, and have small families (1980: 3–79). While Easterlin does not discuss political attitudes, his argument suggests that yuppies (presumably the baby boomers with the greatest ambitions and thus the greatest disappointment) should feel a pessimism and self-absorption which would account both for their rejection of conventional norms and their economic conservatism.[2]

---

[1] At least one source disagrees: according to Price, while baby boomers are embracing traditional values in some areas, they have "a desire for less materialism" (1984: 31). Only baby boomers were interviewed in the survey she reports, however; she presents no data comparing them to their elders.

[2] I do not use the terms *yuppie* and *baby boomer* interchangeably. The latter term is more inclusive, referring to everyone born between 1946 and 1964. The former, while less precise, refers to those in the baby boom cohort who meet some (not usually specified) education and/or income criterion. I refer to baby boomers when citing find-

Alternatively, the combination of lifestyle liberalism and economic conservatism might be accounted for more simply. Academic surveys of voting and public opinion in the 1950s and 1960s commonly found that people of low status were economic liberals for reasons of simple self-interest, while those of high status had education and "sophistication" which made them liberal on other issues, often referred to as issues of "style" (Berelson et al., 1954: 184–85; Campbell et al., 1964: 113–21; Lipset, 1963: 103–04, 318–22). The "style" or noneconomic category encompassed a different set of issues than today—foreign policy and civil rights. (It is striking that as recently as the 1950s, foreign policy and civil rights could be called style issues because they seemed to engage voters' abstract values, not their immediate self-interest!) But it is even more plausible that the well-educated and sophisticated would be liberal on the lifestyle issues which have emerged in recent years (as has been found by Knoke, 1979).

The hypothesized yuppie syndrome would contradict other findings indicating that both youth and education make people more liberal. Stouffer, studying tolerance of nonconformists, argued that education has a liberalizing effect on attitudes and age a conservatizing effect (1955). Davis (1975), replicating Stouffer's study two decades later, likewise finds the young and the well-educated to be most liberal; however, though his younger cohorts were more tolerant than their elders, he finds no evidence for an aging effect and rather argues that the age difference is a cohort effect.

Davis finds the young and well-educated to be more liberal on a variety of attitudes: not just on tolerance of nonconformity but on lifestyle issues as well, and on the General Social Survey federal spending questions, which are clearly related to New Deal liberalism (Davis, 1975 and 1980). These findings would weigh against the hypothesis that yuppies, while liberal on issues of personal freedom, are conservative on economic issues.

Hypotheses

Let us spell out more formally some of the competing hypotheses about the yuppie attitudinal syndrome. These hypotheses distinguish between lifestyle issues (broadly, sexual and gender liberation) and

---

ings (such as Easterlin's) which refer to the entire cohort. I have read no such "findings" on yuppies, only speculations for which the reference group is not clearly identified. In my own analysis below, I do not define who is or is not a yuppie (except in a preliminary analysis); instead, I examine the effects of cohort, education, income, and place of residence separately.

economic issues (social spending and government responsibility for economic welfare).

1. The pure yuppie hypothesis: yuppies (i.e., well-educated, well-off, urban members of the baby boom generation) are more liberal than the rest of the population on issues of lifestyle, more conservative on economic issues. Because this difference is due to cohort size, there are interactions between cohort membership and the variables which determine attitudes of older people. Not only are members of the baby boom cohort and people who live in cities and have higher education and income most likely to manifest economic conservatism and lifestyle liberalism, but the effect of education, income, and urban residence on both sets of attitudes is greater for them.[3]

2. The bifurcation hypothesis: economic conservatism and lifestyle liberalism are determined by high SES. This hypothesis is consistent with the media depiction of the yuppie attitude syndrome. It differs from the "pure" yuppie hypothesis, however, in that it gives no special place to generational membership—yuppies have the attitudes they have not because of any cohort experience, but because education, income, and urban residence have the same effect on their attitudes as on those of the older population.

3. No yuppie difference: the Stouffer/Davis hypothesis predicts that education and youth are associated with liberalism on all issues. Neither cohort size nor coming of age in the 1960s and 1970s has had a unique effect—at least, they have not had the effect widely attributed to them.

The first two hypotheses, while predicting the same pattern of attitudes, predict differences in the causal structures which determine those attitudes. The pure yuppie hypothesis predicts that demographic variables interact to determine attitudes: more education, for example, has a much greater liberalizing effect on the sexual attitudes of the young than it has for the old or middle-aged. The bifurcation hypothesis merely predicts that wealthy, well-educated urbanites differ in attitudes from the rest of the population, but that the determinants of those attitudes are the same.

According to the third hypothesis, the media claim of yuppie conservatism on economic issues is wrong. Regardless of whether the same variables determine the attitudes of yuppies and others, according to this hypothesis, yuppies are not more conservative than the rest of the population on any issues.

---

[3] An interaction would be consistent with the hypothesis that cohort size has an effect. But it would not by itself *demonstrate* any such effect, for it might be due to other characteristics of the cohort, such as the historical circumstances of the period when it came of age.

Methods

The hypotheses were tested on the data of the 1982, 1983, and 1984 NORC General Social Surveys. While the first two were conducted before the yuppies were widely identified as a social group, it is unlikely that their distinctive opinions, if any, were formed in the space of a year or two.

By a loose definition—anyone born after 1945 with a college degree—yuppies make up 8.4 percent of the sample in the 1983 General Social Survey. By criteria closer to those implied by *Newsweek* in its yuppie cover story (1984: 16)—born after 1945 and now earning at least $35,000 in a professional or managerial occupation—they represent a mere 1.1 percent. The present analysis (except its preliminary stage), however, does not rely on any particular definition. Rather, it examines the effects on attitudes of those variables that define the term *yuppie*.

To select attitude variables for analysis, I first examined the zero-order differences in attitude between yuppies (defined as anyone with a bachelor's degree born in 1946 or later) and others on several issues suggested by press accounts. Items for which the difference yielded a gamma greater than .10 were examined further. Scales were formed from several sets of items, most of them related to lifestyle issues: attitudes toward tolerance of pornography, feminism, sexual freedom, nonconformist values for children, and abortion as a method of contraception.[4] Items which formed scales of liberalism on social spending and government intervention also showed substantial difference between yuppies and non-yuppies. (The scaling procedure and the items included in each scale are described in the Appendix.) Yuppies had nearly the same opinions as the general population on questions related to war and militarism, federal spending on social welfare or for cities, and (for the most part) racial issues.

I analyzed the effects on each scale of age, education, income, and urban residence.[5] To increase the sample size, I included the 1982 and/or 1984 General Social Survey samples whenever the same items were

[4] In general, people are far more accepting of abortion if the woman was raped or if the child is likely to have a birth defect than if the woman wants an abortion because she is unmarried, wants no more children, or is poor. I refer to the latter set of reasons as the respondent's attitude to abortion as a method of contraception.

[5] I also examined the effects of occupation, dichotomizing between professionals and others. It had no significant effects, and I have omitted it from the tables even though it is part of the definition of yuppie. Its lack of effect might be due to the fact that while yuppies are called professionals, the occupations most identified with the status are not professional but management occupations; the more likely reason is that the effects of occupational level on attitudes usually turn out to be effects of education and income and disappear when the latter two are controlled.

present. Age was measured in years and then as a pair of dummy variables representing the baby boom cohort (born in 1946 or later, thus including a small number of post-baby boomers) or its immediate predecessor (born between 1933 and 1945). The cohort definition was intended to test whether being in the baby boom cohort has an effect distinct from that of age in general; the second cohort dummy was included to allow for the possibility (though it is not mentioned in the literature) that the yuppies differ from their immediate elders more than from the population as a whole. If age once had a uniform liberalizing effect but no longer does, the middle cohort might be more liberal than either their elders or the baby boom cohort.

Findings

Table 1 presents three regression analyses of each lifestyle issue scale. The first examines the effects of education, family income, urban residence, and two dummy variables for baby boom cohort and middle cohort.[6] The second tests for an interaction between membership in the baby boom generation and income, education, or urban residence.[7] In the final equation age is measured in years, to determine whether the effect of cohort membership could be equally well explained as a linear effect of age.

The pure yuppie hypothesis does not fare very well by these tests. There is little sign either of any interaction effect or even of a distinct generational effect. Consider the equations for tolerance of pornography in Table 1: the first column shows that education and urban residence both affect a respondent's attitudes toward pornography, and that younger people are more accepting than people of middle age, who are in turn more accepting than older people.

The second equation tests for an interaction between membership in the baby boom cohort and income, education, or urban residence. A positive coefficient for an interaction term would mean that education, income, or urban residence had a greater effect on younger people's

---

[6] Income, aggregated in the survey data, was turned into an intervally scaled variable by assigning to each respondent the category midpoint; respondents in the lowest category (below $1000) were arbitrarily assigned $750; those in the highest category ($50,000 and up) were arbitrarily assigned $60,000. To correct for inflation, these values were inversely weighted by the Consumer Price Index for the year before the survey year.

Urban residence was defined dichotomously: all those who lived in a large central city (over 250,000) or its suburbs, or a medium-sized central city (50,000 to 250,000) were classified as urban.

[7] Interactions for the middle-cohort dummy variable were also tested, but none were found worth reporting. Similarly, there were no higher-order interactions of any consequence among the variables included in the regressions reported.

Table 1. Effect of Education, Income, Urban Residence, and Age/Cohort on Lifestyle Issues (Raw Regression Coefficients)

Independent Variable	*Tolerance of Pornography*			*Feminism*			*Sexual Freedom*			*Value Children's Nonconformity*			*Contraceptive Abortion*		
Education	.095*	.094*	.092*	.220*	.228*	.211*	.151*	.166*	.142*	.179*	.166*	.180*	.235*	.161*	.237*
Income	.003	.007	.002	.014*	.020*	.013*	.016*	.025*	.015*	.009*	.010	.009*	.017*	.031*	.016*
Urban	.394*	.404*	.389*	.192	.105	.196	.305*	.456*	.298*	.213	.216	.215	.855*	.919*	.855*
Baby boom cohort	1.793*	2.081*		1.392*	2.294*		1.461*	3.016*		.539*	.153		.274*	−1.295*	
Middle cohort	1.151*	1.123*		.987*	.915*		.858*	.754*		.496*	.510*		.044	.047	
Interaction: baby boom by education		−.005			−.050			−.074			.034			.189*	
Interaction: baby boom by income		−.009			−.016*			−.021*			−.001			−.030*	
Interaction: baby boom by urban		−.013			.219			−.321			−.010			−.163	
Age			−.049*			−.040*			−.043*			−.013*			−.006
Intercept	−1.836*	1.762*	5.106*	2.582*	−2.767*	.118	11.730*	−12.147*	−8.899*	−2.772*	−2.652*	−1.886*	−16.508*	−16.023*	−16.123*
Adjusted R^2	.112	.112	.121	.160	.163	.170	.130	.137	.143	.092	.091	.090	.085	.091	.085
(N)		(2534)			(2058)			(2920)			(1437)			(4099)	

* Coefficient more than twice its standard error.

Table 2. Effect of Education, Income, Urban Residence, and Age/Cohort on Economic Issues, Party Identification, and Liberalism/Conservatism (Raw Regression Coefficients)

Independent Variable	*Social Spending*			*Government Intervention*			*Party Identification*			*Liberal Conservative*		
Education	.042*	.020	.038*	−.124*	−.127*	−.125*	−.083*	−.093*	−.086*	.006	−.008	.007
Income	−.011*	−.009*	−.012*	−.019*	−.019*	−.019*	−.015*	−.018*	−.014*	−.005*	−.004*	−.005*
Urban	.615*	.680*	.614*	.186*	.290*	.186*	.333*	.456*	.330*	.192*	.208*	.193*
Baby boom cohort	1.001*	.467		.457*	.481		.051	−.507		.382*	−.108	
Middle cohort	.630*	.642*		.266*	.268*		.243*	.288*		.234*	.250*	
Interaction: baby boom by education		.060*			.006			.041			.043*	
Interaction: baby boom by income		−.006			.000			.008*			−.001	
Interaction: baby boom by urban		−.153			−.216			−.278*			−.040	
Age			−.028*			−.013*			−.003			−.010*
Intercept	−8.246*	−8.090*	−6.387*	−2.702*	−2.718*	−1.848*	4.643*	4.761*	4.883*	3.615*	3.748*	4.263*
Adjusted R^2	.082	.085	.089	.125	.125	.128	.047	.050	.046	.024	.025	.024
(N)		(3112)			(2575)			(4417)			(3590)	

* Coefficient more than twice its standard error.

tolerance than it did for older people. But all the coefficients are negative (though not statistically significant).

The third equation tests for the effects of age in years. If the model in the first column explains significantly more variance than this model, membership in either the baby boom cohort or the middle cohort exerts a generational effect. The coefficients in the first model show a monotonic increase in tolerance from the oldest to the youngest group. The comparison of the adjusted R^2 for the two models confirms that the effect of age is linear, because the third model has a slightly larger adjusted R^2 (the adjusted R^2 takes into account degrees of freedom, so two adjusted R^2s are comparable even if based on different numbers of independent variables).

Results are similar for feminism and sexual freedom. For both, there is a statistically significant interaction between income and membership in the baby boom. But it is negative, meaning that income has more of a liberalizing effect for older people than for baby boomers.

The analysis of a scale of items reflecting preference for nonconformist behavior in children gives similar results, except for a curvilinear effect of age, but not one which singles out baby boomers. They and the middle cohort are nearly identical, more liberal than the oldest cohort (as the coefficients of the two dummy variables in the first column reveal).

The analysis of attitudes on abortion as a means of contraception, in the last section of Table 1, offers some (mixed) support for the pure yuppie hypothesis. The youngest cohort is more approving than the middle and older cohorts, as the first model shows, and the interaction model shows that it is the best-educated baby boomers who most support contraceptive abortion. The interaction effect for income, however, goes in the opposite direction.

Lifestyle attitudes offer little support for the pure yuppie hypothesis: the lack of interactions (except in the case of contraceptive abortion) suggests that yuppies are no more liberal than can be expected from their age, education, income, and place of residence. Those attitudes do conform to the bifurcation hypothesis, but do not distinguish between it and the hypothesis of no yuppie effect, because those two hypotheses only differ in their predictions regarding economic issues.

Table 2 reports findings regarding economic issues. If the media accounts are correct, yuppies should be opposed to higher federal spending on education, on the environment, and to improve the condition of blacks. But the first section of Table 2 shows that the youngest cohort is considerably more favorable to spending for these purposes than its elders. Though there is a small interaction between baby boom cohort and education, it goes in the wrong direction for the yuppie hypothesis. The third column, in any case, shows that the effect of age is linear.

Spending on education and the environment, if not spending to improve the condition of blacks, might represent a "postindustrial" liberalism (although factor analysis showed that the three variables clustered). The issues analyzed next, however, represent traditional liberalism: agreement that the government should reduce income differentials and that it should solve problems rather than leaving them to the private sector. On government intervention, the well-educated are less liberal than the less well-educated; but it is the youngest cohort which is most liberal.

Table 2 reports the analysis of party identification, with results almost identical (though weaker) to those for government intervention. (A positive score represents identification as a Democrat.) The middle cohort is most Democratic, but the youngest cohort is slightly more Democratic than the oldest. On government intervention and party identification, in contrast to social spending, the better educated (young and old) are more conservative; but the young are not more conservative than the old.

Finally, Table 2 presents the analysis of self-definition as liberal or conservative (an ambiguous variable: people locate themselves on the basis of very different issues (Robinson and Fleishman, 1984), but I present the results because so much of the discussion of yuppie attitudes has been couched in terms of liberalism and conservatism). The multiple correlation coefficient is quite low. But the effects of baby boom cohort and education are positive: yuppies regards themselves as more liberal than the population as a whole, and even than their immediate elders.

Yuppies—A Media Mirage?

These data thus provide almost no evidence for the pattern of yuppie attitudes predicted by both the pure yuppie hypothesis and the bifurcation hypothesis. Unfortunately, the General Social Survey is silent on the attitudes dearest to the yuppies of the media image, physical fitness and consumerism. But while the young and well-educated are liberal on lifestyle issues, their vaunted conservatism on economic issues has failed to materialize.

For a brief period during 1984 and 1985, stories about yuppies and their unique political views filled the press, without much evidence or serious examination. The term apparently first appeared in 1983 in the East Bay *Express* (Kahn, 1983). It became current as a description of the people who voted for Gary Hart in the 1984 primaries, especially in New Hampshire and Massachusetts. According to Dotty Lynch, Hart's chief pollster, news reports attributed the term to the Hart campaign staff, but she first heard it when she read those reports (personal communication). Media attention to the yuppies was short-lived: by mid-1985, journalists (including the writer who coined the term) were declaring the yuppie phenomenon over (Kahn, 1985; cf. Dowd, 1985). None of the obituaries, however, questioned the accuracy of the earlier accounts.

How can we explain the short life, untimely death, and general inaccuracy of the media's image of the yuppie? The yuppies' reputation for conservatism may have been due to the absence of the radical politics so visible among the youth (perhaps some of today's yuppies themselves) of a decade or two ago. But I suspect that a more important factor was at work: the image did not receive careful scrutiny because it plausibly explained a lot of observations and because it provided an opportunity to practice common journalistic routines.

Yuppies came of age in imposing numbers. They visibly transformed the American cityscape with health clubs, gourmet delis and remodeled houses. In the process they became a marketing target, and since newspapers, magazines, and television are marketed, journalists had to pay attention to them. At the same time, another trend appeared to be occurring which demanded explanation: a rightward drift in public opinion.[8] It was easy to believe that the yuppies were responsible, as they practiced a conspicuous consumption often associated with conservative opinions.

Post hoc ergo propter hoc. But the yuppie story not only seemed to explain a conservative trend; it fit the patterns of journalistic storytelling. Journalists characteristically structure stories around events rather than issues (Tuchman, 1978). The arrival of the yuppies qualified as an event; it could be dramatized and visualized. With the image of the yuppie, the conservative drift could be personalized and precisely located on Columbus Avenue or in Back Bay.

---

[8] The extent to which public opinion is becoming more conservative is itself a subject of controversy, but one beyond the scope of this article. Some argue that the trend is strong; others see a waning, but not a reversal of liberalism. Still others argue that the election and reelection of Ronald Reagan create the appearance of a conservative tide, but that his electoral success depends not on ideological conservatism but on other short-term factors. See, e.g., Davis, 1980; Himmelstein and McRae, 1984; Lipset, 1985; Miller and Shanks, 1982; *Public Opinion*, 1984; Russett and DeLuca, 1981; Smith, 1981.

If public opinion is in fact becoming more conservative, it may be that the yuppies are moving in the same direction as the whole population, but no farther or more quickly. Davis (1980) argues that a conservative shift can be detected, but that—up to 1978 at least—it was no more pronounced in one demographic group than any other. The evidence presented in this article does not suggest any conservative trend among the young. The youngest cohort is the least conservative on two of the attitudes analyzed in Table 2, and less conservative than the oldest generation on the third (party identification).

If the media image of yuppies offered an intellectually satisfying answer to some difficult questions, the press's infatuation with that image was probably due as well to a social process common among journalists. Newspeople are prone to "pack journalism," especially when they do not fully understand something: "stories spread horizontally, from news organization to news organization. . . . A critical mass of attention heaps up; then [a story] 'cannot be ignored' " (Gitlin, 1980: 99–100). Its currency elsewhere in the media gives it a veneer of "objectivity" which legitimizes its presence in a reporter's story.

It was easy to find isolated examples apparently supporting the claims made about yuppies, but normal journalistic routines do not lend themselves to testing a claim about the opinions of an ill-defined social group, and journalists are not very well versed in the social science methodology which would enable them to verify it (Singer and Endreny, 1986). If most journalists do not have those tools, the public has them even less. When we are puzzled by something new, we select from the many experiences available to us those that apparently provide the best explanation; and among those experiences are the media reports themselves. When an explanation is repeated in the media as often as the yuppie explanation was, it takes on a self-fulfilling character; news accounts corroborate each other and legitimize the story in the minds of readers and viewers, who then recognize some of its elements in their own experience. A member of the media-attentive public, strolling down the street and spotting a jogger, can reflect, "Yes, the country is becoming more conservative."

The belief may have been self-confirming for politicians and lawmakers as well, many of whom are eager to cater to what they perceive as demands for cutbacks in government services and economic assistance. Today's young people are not necessarily opposed to such government programs. They want no restrictions on their lifestyles, but they have not totally abandoned the social concerns that inspired the young in the past.

The rapid spread of the yuppie story and its equally rapid disappearance are themselves grounds for suspecting that its diffusion was due to a social process rather than a process of verification. No social group adopts, or abandons, a common culture overnight. The yuppies were not born yesterday, after all—some have reached 40, and anything that is said about them, however recently discovered, has probably been true (or false) for longer than most journalists have been aware of it.

Appendix: Scale Construction

Each scale is the sum of the standard scores of its items, each item weighted so that a positive score represents liberalism. The survey years in which all the items for each scale (and the single-item indicators) appear, and each item's mnemonic and abbreviated question wording, are listed below.

TOLERANCE OF PORNOGRAPHY (1983, 1984)

PORNLAW	Laws against distribution of pornography
PORNMORL	Pornography leads to breakdown in morals
PORNRAPE	Pornography leads to rape
XMOVIE	Seen X-rated movie in last year

FEMINISM (1983, 1984)

FEHOME	Women take care of home not country
FEPOL	Women not suited for politics
FEPRES	Vote for woman president
FEWORK	Should women work

SEXUAL FREEDOM (1982, 1983)

PILL	Birth control information should be available
TEENPILL	Birth control info available to teens
SEXEDUC	Sex education in public schools
PREMARSX	Sex before marriage

VALUES FOR CHILDREN (1983)

ROLE	Acts like boy or girl
OBEYS	Obeys parents
CONSIDER	Considerate of others
INTEREST	Interested how and why things happen

ACCEPT CONTRACEPTIVE ABORTION (1982, 1983, 1984)

ABPOOR	If can't afford more children
ABSINGLE	If not married
ABNOMORE	Married—wants no more children
ABANY	If woman wants for any reason

SOCIAL SPENDING LIBERALISM (1982, 1983, 1984)

NATEDUC	Federal spending for education
NATENVIR	Federal spending for environment
NATRACE	Federal spending for condition of blacks

GOVERNMENT INTERVENTION (1983)

EQWLTH	Govt should reduce income differences
HELPNOT	Govt solve problems, not leave to individuals or business
PARTYID	Party identification (1982, 1983, 1984)
POLVIEWS	Liberal/conservative (1982, 1983, 1984)

References

Abramson, Paul R.
1983 Political Attitudes in America: Formation and Change. San Francisco: W. H. Freeman.
Bellah, Robert N. et al.
1985 Habits of the Heart: Individualism and Commitment in American Life. Berkeley: University of California Press.
Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee
1954 Voting: A Study of Opinion Formation in a Presidential Campaign. Chicago: University of Chicago Press.
Boaz, David
1985 "In '88, who'll win the baby boomers?" New York Times, November 7, A35.
Campbell, Angus et al.
1964 The American Voter: An Abridgment. New York: John Wiley.
Davis, James A.
1975 "Communism, conformity, cohorts, and categories: American tolerance in 1954 and 1972–73." American Journal of Sociology 81:491–513.
1980 "Conservative weather in a liberalizing climate: change in selected NORC General Social Survey items, 1972–78." Social Forces 58:1129–56.
Dowd, Maureen
1985 "Retreat of the yuppies: the tide now turns amid 'guilt' and 'denial.' " New York Times, June 28, B1, B4.
Easterlin, Richard A.
1980 Birth and Fortune: The Impact of Numbers on Personal Welfare. New York: Basic Books.
Gitlin, Todd
1980 The Whole World Is Watching: Mass Media in the Making and Unmaking of the New Left. Berkeley: University of California Press.
Himmelstein, Jerome L., and James A. McRae, Jr.
1984 "Social conservatism, new Republicanism, and the 1980 elections." Public Opinion Quarterly 48:592–605.
Inglehart, Ronald
1977 The Silent Revolution. Princeton: Princeton University Press.
Jennings, M. Kent, and Richard G. Niemi
1981 Generations and Politics: A Panel Study of Young Adults and Their Parents. Princeton: Princeton University Press.
Jones, Landon Y.
1981 Great Expectations: America and the Baby Boom Generation. New York: Ballantine Books.
Kahn, Alice
1983 "Yuppie!" Express: The East Bay's Free Weekly 5 (June 10), 1, 6–7.
1985 "Where have all the yuppies gone?" Express: The East Bay's Free Weekly 7 (July 26), 1, 8–12
Knoke, David
1979 "Stratification and the dimensions of American political orientations." American Journal of Political Science 23:772–91.
Lipset, Seymour Martin
1963 Political Man: The Social Bases of Politics. Garden City: Doubleday Anchor.
Mannheim, Karl
1969 "The problem of generations." Pp. 352–69 in Barry McLaughlin (ed.), Studies in Social Movements: A Social Psychological Perspective. New York: Free Press.
Miller, Warren E., and J. Merrill Shanks
1982 "Policy directions and presidential leadership: alternative interpretations of the 1980 presidential election." British Journal of Political Science 12:299–356.

Newsweek
1984 "The year of the yuppie." Newsweek (December 31):14–31.
Price, Barbara A.
1984 "What the baby boom believes." American Demographics 6 (May): 30–33.
PS
1985 "The elections, the economy, and public opinion: 1984." PS 18 (Winter), 28–38.
Public Opinion
1984 "Values: generations apart." Public Opinion (December/January): 21–37.
Robinson, John P., and John A. Fleishman
1984 "Ideological trends in American public opinion." Annals of the American Academy of Political Science 472 (March): 50–60.
Russett, Bruce, and Donald R. DeLuca
1981 " 'Don't tread on me': public opinion and foreign policy in the eighties." Political Science Quarterly 96:381–99.
Singer, Eleanor, and Phyllis Endreny
1986 "The reporting of social science research in the mass media." In Sandra Ball-Rokeach and Muriel Cantor (eds.), Media, Audience, and Social Structure. Beverly Hills: Sage Publications.
Smith, Tom W.
1981 "General liberalism and social change in post World War II America: a summary of trends." Social Indicators Research 10:1–28.
Stouffer, Samuel
1955 Communism, Conformity, and Civil Liberties. New York: Doubleday.
Tuchman, Gaye
1978 Making News: A Study in the Construction of Reality. New York: Free Press.
Turner, Ralph H.
1976 "The real self: from institution to impulse." American Journal of Sociology 81:989–1016.
Veroff, Joseph, Elizabeth Douvan, and Richard A. Kulka
1961 The Inner American: A Self-Portrait from 1957 to 1976. New York: Basic Books.
Yankelovich, Daniel
1981 New Rules: Searching for Self-Fulfillment in a World Turned Upside Down. New York: Random House.

Do Yuppies Matter? Competing Explanations of Their Political Distinctiveness

MICHAEL DELLI CARPINI AND LEE SIGELMAN

IN FUTURE years, 1984 may be viewed as a landmark in the changing political landscape of the United States, notable primarily for the widening gap between men and women on certain political issues and for the exodus of college students, blue-collar workers, and white southerners from the Democratic coalition. In 1984, however, none of these developments received anything approaching the media attention that was lavished on the emergence of young urban professionals, or "yuppies," as a political force. Indeed, 1984 was, as *Newsweek* magazine proclaimed in a cover story, "The Year of the Yuppie," and yuppies were held largely responsible for the year's most surprising political development, the strong primary showing of Gary Hart, himself a yuppie emeritus.

Three separate but related ideas lie beneath the surface of the much-ballyhooed yuppie phenomenon. First, young urban professionals are thought to comprise a distinct "class." That is, a unique confluence of demographic forces—youth, urban residence, and professional

---

Abstract "Yuppies"—young urban professionals—have often been portrayed as an important new political force. Using data from the 1972–1985 cumulative General Social Survey data file, this article tests various explanations of the political distinctiveness of yuppies. The findings call into question the assumption that youth, urbanity, and professional status combine to form a distinct yuppie political culture above and beyond what would be expected on the basis of an additive combination of these factors. For this and other reasons, the political importance of the yuppie phenomenon appears to have been overstated.

Michael Delli Carpini is Assistant Professor in the Department of Political Science, Rutgers University. Lee Sigelman is Professor in the Department of Political Science, University of Kentucky.

Public Opinion Quarterly Vol. 50:502–518 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-502/$2.50

status—is thought to distinguish yuppies socially and culturally from other young people, from other city dwellers, from other professionals, and, of course, from the nonyoung, the nonurban, and the nonprofessionals as well. Second, stemming from this unique demographic profile and its social and cultural offshoots are said to be a distinctive set of political attitudes and opinions. Yuppies are usually (though not always) seen as embodying the "new politics" of the late 1960s and early 1970s (Miller and Levitin, 1976) *and* the economic conservatism of the 1980s. Third, as a consequence of their distinctive political views, yuppies are considered to be an increasingly significant voting bloc; that is, they are supposed to require special appeals from candidates and parties, and respond to such appeals in much the same way as do blacks, union members, and other groups in American society—by supporting the parties and candidates who appeal directly to their group interests.

This study probes the validity of these ideas. Using data from the 1972–1985 NORC General Social Surveys (GSS), we begin by addressing a pair of key descriptive issues: How significant a part of the American electorate are yuppies? And how distinctive are their political attitudes and opinions? We then turn to a more analytical question: To what can the political distinctiveness of yuppies be attributed? Is it primarily their youth? Their urban residence? Their professional status? The sum total of all three? Or is the whole somehow distinct from the assemblage of parts? Drawing on prior studies of the "new class" and of the relationships between political orientations and age, class, and residence, we outline and test three different interpretations of the roots of the yuppie political phenomenon.[1]

Describing the Yuppie Phenomenon

How large a segment of the American public do yuppies comprise? How distinctive are they politically? Before we can answer these questions, we must address a prior issue: What is a yuppie? With tongue in cheek, Piesman and Hartley define a yuppie as a person who: (1) resides in or near a major city; (2) claims to be between the ages of 25 and 45; (3) lives on aspirations of glory, prestige, recognition, fame, social status, power, money, or any of the above; and (4) brunches on the weekend or works out after work (1984: 12). This definition, which is more clever than most but representative of many, employs both demographic and lifestyle criteria and thereby confuses the alleged consequences of the yuppie phenomenon, understood demographically, with the phenomenon itself. Our own approach is to construe "yuppie" literally: a yuppie is someone who is young, who resides in an urban setting, and who works in a professional occupation. This approach hardly resolves all the important measurement issues (What is young? What is urban? What is a profession?), but it does permit greater discrimination than do many previous usages of the term, which has sometimes been stretched to include all young professionals, urban or not, or even all young people who hold college degrees. Groups defined as broadly as that may merit attention, but calling them yuppies only confuses matters.[2]

---

[1] Gopoian et al. (1985) have recently used data from the GSS to test Ladd's (1978, 1979) old class/new class interpretation of conflict within the Democratic Party, a topic somewhat related to but quite different from our own interest.

For our purposes, *young* is defined as being between 18 and 39 years of age, *urban* as living in an urban area of at least 100,000 people, and *professional* as working in a job classified as professional or technical by the U.S. Census Bureau.[3] According to the pooled 1972–85 GSS data, young people, urbanites, and professionals all represent sizable proportions of the adult population: 46 percent of the GSS respondents are 39 or younger, 26 percent reside in urban areas, and 15 percent work in professional occupations.[4] However, young urban professionals—people who combine all three of the defining criteria of the yuppie—are quite rare. Yuppies, as defined here, comprise only 2.4 percent of the 1972–85 GSS respondent pool, only 5 percent of those less than 40 years of age, only 9 percent of the urbanites, and only 16 percent of the professionals. In short, yuppies are an extremely small proportion of both the population as a whole and of the relevant subpopulations.[5]

How distinctive are the political attitudes and opinions of this small but highly publicized group? In Tables 1 through 3, we compare yuppies to the rest of the adult public on several key political dimensions. The responses summarized in Table 1 make it clear that there are some

---

[2] While we use a fairly strict definition of yuppies, our exploratory analyses using other, broader definitions (such as holding a college degree rather than being a professional) suggest that the conclusions reported below would apply to other definitions of yuppies as well.

[3] The cutoff of 39 years old is somewhat arbitrary, and has occasioned much bitterness on the part of one of the co-authors. Replicating the analysis with younger cutoffs (to 35 years old) leads to essentially the same conclusions as those drawn in the text.

[4] The pooled GSS data set is used because it provides a large enough number of cases (up to 16,572) to allow for an examination of yuppies (404 cases). Replicating the analysis on a survey-by-survey basis suggests that no underlying trends from 1972 to 1985 are hidden by our reliance on the cumulative data set.

[5] The 2.4 percent figure is consistent with the estimates of Markus and Jennings (1985), who used panel data, and Bruce-Briggs (1979), who used census data. This figure is quite stable over the 15 years of the GSS.

Table 1. Differences Between Yuppies and Nonyuppies on Selected Political Opinions

	Yuppies	*Nonyuppies*	*Sig.*
Independent	44%	34%	***
Liberal	49	27	***
Vote Democratic	63	45	***
Taxes too high	70	69	NS
Favor capital punishment	56	69	***
Favor gun control	83	72	***
Courts too harsh	8	4	***
Approve of wiretaps	13	19	**
Legalize marijuana	55	23	***
U.S. stay in U.N.	92	78	***
Communism worst form of government	28	55	***

* = $p < .05$; ** = $p < .01$; *** = $p \leq .001$; NS = nonsignificant.

substantial differences between the political opinions of yuppies and other Americans. Yuppies are less likely to identify with a political party but, by a large margin, more likely to call themselves liberals and to vote Democratic in presidential elections. On particular political issues, yuppies are generally more liberal as well: they are more opposed to capital punishment and wiretaps, disapprove less of communism, and are more likely to favor gun control, to consider the courts too harsh, to support the legalization of marijuana, and to advocate continued American membership in the United Nations. However, yuppies do not stand out from other Americans with respect to dissatisfaction with the taxes they pay.

Table 2 summarizes some additional evidence concerning the political distinctiveness of yuppies, in the form of the Stouffer (1955) battery of questions concerning tolerance for those who hold unconventional views or practice unconventional lifestyles. In every instance—no matter whether the target group is composed of atheists, socialists, racists, communists, militarists, or homosexuals, and no matter which civil liberty is in question—yuppies express greater tolerance than do other Americans. Not only are these differences statistically significant; many of them are very appreciable as well, with yuppies in some cases being twice as likely as nonyuppies to adopt the tolerant position.

Part of the popular image of yuppies is that they are socially liberal but economically somewhat conservative, and the large percentage of yuppies who feel that taxes are too high (see Table 1) is consistent with that image. In Table 3 we pursue this possibility by examining answers to a series of questions that ask whether the amount the federal govern-

Table 2. Differences in Political Tolerance Between Yuppies and Nonyuppies (percentage that would allow each group to speak, teach college, or have a book in the library)

	Yuppies	*Nonyuppies*	*Sig.*
Atheists			
Speak	91%	66%	***
Teach	83	44	***
Book	91	62	***
Socialists			
Speak	94	80	***
Teach	91	60	***
Book	94	73	***
Racists			
Speak	79	60	***
Teach	67	43	***
Book	84	63	***
Communists			
Speak	88	58	***
Teach	79	43	***
Book	88	59	***
Military rule			
Speak	86	55	***
Teach	76	39	***
Book	86	58	***
Homosexual			
Speak	93	67	***
Teach	90	56	***
Book	86	58	***

* = $p < .05$; ** = $p < .01$; *** = $p < .001$.

ment spends in several policy areas is "too much," "too little," or "about right." Playing against their stereotype, yuppies stand out for their tendency to say that too *little* is being spent in most areas: on the space program, the environment, national health, the cities, education, black welfare, and social welfare. On the other side of the coin, yuppies are unusually opposed to defense spending, but this opposition seems more indicative of liberal leanings than of a proclivity toward economic conservatism. The lack of any significant difference between yuppies and other Americans on spending for foreign aid and for combating crime and drug use reflects, in large measure, the noncontroversial nature of these programs: very few Amercians, whatever their political outlook, oppose stepped up anticrime and antidrug abuse programs, and very few favor increased foreign aid. The fact that yuppies are indistinguishable from other Americans in their opposition to programs designed to equalize wealth is more interesting, for it serves to emphasize that even though yuppies are apparently more liberal than other Americans, their leftist leanings stop well short of opposition to the basic tenets of individual enterprise.

Table 3. Difference Between Yuppies and Nonyuppies on Attitudes About Government Spending

Too Little Spent on:	*Yuppies*	*Nonyuppies*	*Sig.*
Space program	21%	12%	***
Environment	79	57	***
National health	74	61	***
Cities	66	50	***
Crime	68	70	NS
Drug use	57	63	NS
Education	72	55	***
Black welfare	47	31	***
Defense	11	27	***
Foreign aid	5	4	NS
Welfare	27	19	**
Equalize wealth	29	30	NS

* = $p < .05$; ** = $p < .01$; *** = $p < .001$; NS = nonsignificant

Dissecting the Yuppie Phenomenon

It is one thing to document political differences between yuppies and other Americans, but it is something else to pinpoint the sources of those differences. Is the political distinctiveness of yuppies primarily attributable to their youth, their urban residence, or their professional status? Is it the cumulative effect of these factors that makes yuppies what they are politically? Or do youth, urban residence, and professional status *in combination* have attitudinal effects above and beyond their cumulative effects?

In dissecting the yuppie phenomenon, we shall pursue three different sets of answers to these questions. To begin with the most complex interpretation, which we term the *yuppie class thesis*, the yuppies' youth, urban residence, and professional status interact in a complex way to produce a culture that nurtures a certain social and political outlook. This interpretation is a variant of the "new class" arguments of Bell (1973, 1979), Ehrenreich and Ehrenreich (1977), Gouldner (1979), Kristol (1972), Ladd (1978), and others, and it also has roots in Reich's (1970) "greening of America" and Inglehart's (1977) "postmaterialism" arguments. All these ideas revolve around the development of a new class or subclass that is (depending on which specific argument one consults) young (Inglehart; Podhoretz, 1979), professional (Ehrenreich and Ehrenreich) or upper-middle class (Wildavsky, 1979) or well educated (Ladd), and urbane if not urban (Bruce-Briggs, 1979). Whether the political outlook of this new class is liberal, conservative, or some new ideological hybrid is a point of contention (see Harrington, 1979; Markus and Jennings, 1985), but on balance the arguments point more toward a type of liberalism than conservatism, a conclusion strongly supported by the data summarized in Tables 1 through 3 above.

There is a second way, which we term the *traditional groups thesis*, to approach the "new class" idea in general and the yuppie phenomenon in particular. Rather than viewing the political distinctiveness of yuppies as an outgrowth of the *interaction* of youth, urban residence, and professional status, which combine in some unique way, it is possible (as well as theoretically more parsimonious) that yuppies are distinctive because they are, in effect, the sum of their parts: they are young, and therefore partake of the political attitudes of the young; they live in cities, and therefore share the political outlooks of urban dwellers; and they are professionals, and therefore think, socially and politically, like other professionals. From this perspective, yuppies do not constitute a distinct class, and no unique fusion of age, residence, and professional status is necessary to explain their distinctive social and political orientations.

Finally, a variant of the traditional groups interpretation, the *special effects thesis*, suggests that age, residence, and status come to the surface politically in rather different ways in different situations. From this perspective, being a yuppie is less like being black or a member of the middle class than it is like being a middle-class black: one is likely to express oneself differently at different times, depending on the circumstances in which one finds oneself; sometimes class considerations predominate, but at other times racial ones do (Bennett, 1980). Thus, for example, yuppies may exhibit greater liberalism and partisan independence as a reflection of their youth (Delli Carpini, 1986), more tolerance as a reflection of their urban residence (Stouffer, 1955), and more opposition to certain government programs as a consequence of their professional status (Brint, 1984). The idea here is not so much that yuppies are, politically, the sum of the parts; rather, a different part may come to the fore politically on each set of issues.

In distinguishing among the joint *interactive* effects of youth, urban residence, and professional status (the yuppie class thesis), their cumulative or *additive* effects (the traditional group thesis), and their differentiated *individual* effects (the special effects thesis), we are not engaging in esoteric hair-splitting. Rather, we are outlining three fundamentally different accounts of the yuppie phenomenon. The yuppie class thesis argues that these separate, sometimes competing forces have resolved themselves into a distinct ideology. This ideology might not fit neatly into a traditional liberal or conservative definition, but it can be identified, and once identified can be appealed to and expressed consistently over time. The traditional groups thesis argues that the important groups to focus on are not yuppies but the larger and more traditional categories defined by age, occupation, and/or place of residence. From this perspective, there is really nothing politically distinctive about yuppies, who can be looked upon simply as the sum product of various demographic categories. The special effects thesis leads in the same general direction, but not entirely. Here the political distinctiveness of yuppies is viewed on an issue-by-issue basis rather than across the board: particular components of the distinctive yuppie political outlook are seen as stemming from particular demographic elements in the yuppie profile. In sum, whereas the yuppie class thesis sees a gestalt, an overall configuration that is distinct from its individual components, the traditional groups and special effects theses (the latter far more than the former) see a hodgepodge of disparate elements, each with its own peculiar effects.

How can we determine which of these interpretations is the most accurate? The question boils down to whether the political distinctiveness of yuppies is the result of the separate effects of age, residence, and profession, rather than the interaction of these characteristics. If it is, then the secondary question becomes whether youth, urban residence, and professional status have more or less equivalent attitudinal effects, or whether each comes to the fore on certain issues and is relatively ineffectual on others. Viewed in this light, what we have is a relatively straightforward problem in the analysis of variance. We can simply categorize the pooled GSS respondents on the basis of the three variables of interest (39 or younger vs. over 39 years of age, urban vs. nonurban residence, and professional vs. nonprofessional occupation), and then examine the main, two-way, and three-way interaction effects of these characteristics on the political attitudes and opinions we have already examined. To avoid confounding the results, we also control for race and sex.[6] If the yuppie class thesis is correct, then the three-way interactions of age, residence, and profession should be substantial. If the traditional groups or the special effects thesis is correct, then the main effects of youth, urban residence, and professional status should predominate. If the relative influence of these factors varies a good deal from one item to the next, then the special effects thesis would seem applicable, but if these influences are reasonably consistent we would have evidence on behalf of the traditional groups thesis.

[6] The results presented below are from a classical analysis of variance model, in which effects are adjusted first for the covariates (controls), then for the main effects, and finally for the interaction effects. This approach was selected because it most closely parallels the theoretical model underlying the research. In point of fact, however, the use of a regression approach, in which all three levels of effects are controlled for simultaneously, does not change the substantive conclusions we draw, though it does decrease the number of significant main effects slightly.

Table 4. Main, Two-Way, and Three-Way Interaction Effects of Being Young, Urban, and Professional on Selected Opinions (Controlling for Race and Sex)

	Main Effects			*2-Way Interactions*			*3-Way Interactions*
	Young	*Urban*	*Prof.*	*Y-U*	*Y-P*	*U-P*	*Y-U-P*
Party ID	NS	.05**	−.06**	NS	.03**	.02	NS
Ideology	.06**	.04**	.01	NS	.02	NS	NS
Vote choice	.03*	.09**	−.04*	.01	.03**	.08	NS
Taxes	NS	NS	NS	NS	NS	NS	NS
Capital pun.	.02	.03**	.06**	.03**	NS	NS	NS
Gun control	.02*	.08**	.05**	NS	.03**	NS	NS
Courts	.05**	.01*	.01*	.02*	NS	.02	.01
Wiretaps	.05**	.02	−.03*	NS	NS	.04*	NS
Marijuana	.19**	.08**	.08**	.02**	NS	.06**	NS
UN	.09**	NS	.07**	NS	NS	NS	NS
Communism	.06**	.07**	.04**	.01	NS	NS	NS
Average	.05	.04	.02	.01	.01	.02	.00
	(.08)	(.05)	(.04)				

NOTE: All entries are *net* deviations from the mean on scales with a range of one. A positive entry represents a more liberal score. Only deviations with $p < .05$ are entered. * = $p < .01$; ** = $p < .001$; NS = nonsignificant. Numbers in parentheses are average betas adjusted for independent variables and covariates.

Table 4 presents the analysis of variance results for the set of issues originally analyzed in Table 1.[7] The cell entries are the differences in the mean scores between the group in question and the rest of the pooled sample, with race and sex held constant. To aid in interpretation of the results, all variables have been coded to have a possible range of one, and a positive entry indicates a more liberal score. For example, the .05 entry for urbanites on party identification means that, controlling for age, occupational status, sex, and race, the mean score of urbanites is 5 percent more Democratic than that of those who do not live in urban areas. Because of the difficulty in interpreting statistically insignificant deviations from the mean (especially for interaction effects), only entries that are significant at the .05 level are included in this and the ensuing tables. We do, however, on occasion allude to these nonsignificant deviations when they appear suggestive.

---

[7] Since Multiple Classification Analysis does not calculate deviations from the mean for interaction effects, these were estimated separately. This was done by replicating the ANOVA models, but treating significant interaction terms from the first model as main effects in the second; for example, being a young professional became a main effects category in the second-round ANOVA. The deviation from the mean of this combined effect was then compared to the deviation from the mean that would result from the simple additive effects of the characteristics in question, all other controls being the same. The difference between the expected and the actual deviations is the contribution of the interaction effect to the overall score. To return to the example, if the first ANOVA found that there was a significant interaction effect between age and occupation, a second ANOVA was run, identical to the first except that age and occupation were combined to form a single main effects category (young and professional or not young and professional). All other factors and covariates remained the same. The deviation from the mean for young professionals was then compared to the summed deviation of being young and of being professional in the first model. The difference between these totals is the interactive effect of the two characteristics, controlling for place of residence, race, and sex.

As can be seen by examining the three-way interactions, being a yuppie contributes practically nothing to the political stands of this segment of the population. Only attitudes toward harsher court sentencing are significantly affected by being young *and* urban *and* professional, leading to a 1 percent more liberal stance than would otherwise be expected. On this first set of items, then, the yuppie class thesis does not fare at all well in accounting for the political distinctiveness of yuppies.

There is somewhat more evidence that the two-way interactions contribute to a distinct and liberal political profile. Being young *and* urban leads to slightly more pronounced Democratic vote choice and to more liberal views concerning capital punishment, court sentencing, the legalization of marijuana, and communism as a form of government. Being young *and* a professional leads to slightly more Democratic partisanship and vote choice, a more liberal self-image, and greater support for gun control. And being urban *and* professional leads to more Democratic identification and vote choice, and to more liberal stands on sentencing, wiretapping, and the legalization of marijuana. Again, however, most of the two-way interactions (58 percent) are not significant, and those that are significant tend to be of modest magnitude.

It is in the main effects of age, residence, and occupation that by far the greatest political impacts are found. Being young is associated with more liberal stands on all opinions except party identification and concern over taxes, neither of which is significantly related to age. Urbanites are more liberal on every issue except taxes and American membership in the United Nations, both of which are unrelated to place of residence. Professional status is significantly related to positions on all issues except taxes, but these relationships vary in direction. When we control for the other variables in the model, professionals are slightly more likely to call themselves liberals and to take the liberal view of capital punishment, gun control, court sentencing, the legalization of marijuana, the United Nations, and communism. They are, however, more oriented toward the Republican Party (in both identification and vote) and are more likely to take a permissive attitude toward wiretapping. This now-liberal, now-conservative mixture in the political outlook of professionals bears out, to some extent at least, arguments made by both sides in the debate discussed earlier. What is interesting, however, is that for each issue on which professionals as a group are more conservative than the norm, there is a significant *liberal* interaction for young professionals and/or urban professionals. On these issues, then, the main, conservatizing effect of professional status is offset by a more complex effect in which professional status is involved.

The bottom row of Table 4 shows the mean deviation for each of the main and interaction effects across all eleven items. While none of these effects is truly dramatic, the effects that do emerge are clearly the result of main effects. Almost 75 percent of the liberal tendencies of young urban professionals result from the separate main effects of these three characteristics. The three-way interactions are essentially nonexistent. The two-way interactions are somewhat more important than the three-way, and when they do come into play it is always to increase liberal tendencies. Although one must be extremely cautious in interpreting nonsignificant interactions, this pattern is borne out there as well (not shown).

Table 5 summarizes the analysis of variance results for the political tolerance items. The same pattern of findings that we noted in Table 4 recurs in Table 5: almost no three-way effects, some two-way effects, and dominant main effects. There are a few differences, however. In the case of political tolerance, all but one of the significant two-way interactions are between being young and professional. In addition, in a complete reversal of the findings in Table 4, all of these interactions have a conservatizing effect: young professionals are less tolerant than their occupation and their age would predict. The main effects of age and occupational status are the most important sources of tolerance for yuppies, with place of residence being considerably less important. Professionalism does not demonstrate the same inconsistent effects on political tolerance that it does for general issue stands. Most of all, the main effects are dominant, and once again they point in the direction of greater liberalism. More than 85 percent of the impact of being a yuppie is the result of being young and of being professional, and over 95 percent of the impact is due to main effects. Only one (out of a possible 18) significant effects can be attributed to the three-way interaction among youth, urban residence, and professional status, and like the two-way interactions between age and occupation, this effect is to decrease the level of tolerance. The nonsignificant three-way interactions also tend to decrease tolerance (by an average of .02), but again one can have little confidence in these findings. The nonsignificant two-way interactions are small and erratic.

Finally, Table 6 shows the analysis of variance results for opinions concerning government spending. The overall pattern is similar in most

Table 5. Main, Two-Way, and Three-Way Interaction Effects of Being Young, Urban, and Professional on Political Tolerance (Controlling for Race and Sex)

	Main Effects			*2-Way Interactions*			*3-Way Interactions*
	Young	*Urban*	*Prof.*	*Y-U*	*Y-P*	*U-P*	*Y-U-P*
Atheists							
Speak	.23**	.05**	.19**	NS	−.03**	NS	NS
Teach	.30**	.07**	.21**	NS	NS	NS	NS
Book	.22**	.07**	.21**	NS	−.05**	NS	NS
Socialists							
Speak	.15**	.04*	.12**	NS	−.04	NS	NS
Teach	.22**	.06**	.14**	NS	NS	NS	NS
Book	.16**	.04	.15**	NS	NS	NS	NS
Racists							
Speak	.13**	.04*	.15**	NS	NS	NS	NS
Teach	.15**	.04*	.14**	NS	NS	.05	NS
Book	.11**	.04*	.18**	NS	−.03*	NS	NS
Communists							
Speak	.19**	.08**	.21**	NS	−.02*	NS	NS
Teach	.22**	.08**	.22**	NS	NS	NS	NS
Book	.21**	.08**	.22**	NS	−.05**	NS	NS
Military Rule							
Speak	.22**	.04*	.21**	NS	NS	NS	NS
Teach	.26**	.04**	.20**	NS	NS	NS	−.03
Book	.22**	.05**	.22**	NS	−.04*	NS	NS
Homosexuals							
Speak	.17**	.08**	.20**	NS	−.03*	NS	NS
Teach	.24**	.09**	.22**	NS	NS	NS	NS
Book	.21**	.08**	.20**	NS	−.03**	NS	NS
Average	.20	.06	.19	.00	−.02	.00	.00
	(.21)	(.06)	(.14)				

NOTE: All entries are *net* deviations from the mean on scales with a range of one. Positive entries indicate greater tolerance. Only deviations with $p < .05$ are entered. * = $p < .01$; ** = $p < .001$; NS = nonsignificant. Numbers in parentheses are average betas adjusted for independent variables and covariates.

respects to those found in Tables 4 and 5, except that the interaction effects in Table 6 are even more conspicuously absent. Once again, the three-way interactions are nonsignificant, both statistically and substantively. Thirty of the 36 two-way interactions are also nonsignificant. Of the six significant interactions, four are in the direction of slightly more conservative stands among young urbanites, young professionals, or urban professionals. The nonsignificant two-way interactions again show no discernible pattern.

Again it is the main effects that are most important elements in shaping the political distinctiveness of yuppies. Young people favor government spending more than older people do on every issue except crime, drug use, and equalizing wealth, where there are no significant age-based differences, and defense, where the young favor less spending. Since favoring less spending on defense is the more liberal opin-

Table 6. Main, Two-Way, and Three-Way Interaction Effects of Being Young, Urban, and Professional on Opinions Concerning Government Spending (Controlling for Race and Sex)

Too Little Spent on:	*Main Effects*			*2-Way Interactions*			*3-Way Interactions*
	Young	*Urban*	*Prof.*	*Y-U*	*Y-P*	*U-P*	*Y-U-P*
Space	.07**	.03*	.11**	NS	NS	−.03	NS
Environment	.15**	.05**	.05**	NS	NS	NS	NS
Health	.04**	.01	NS	NS	NS	NS	NS
City	.08**	.08**	.04**	−.02*	NS	NS	NS
Crime	NS	.01*	NS	.01**	NS	NS	NS
Drug use	NS	−.01	−.02	NS	NS	NS	NS
Education	.11**	.03**	.03**	NS	NS	NS	NS
Blacks	.06**	.03**	.06**	NS	NS	NS	NS
Defense	−.08**	−.04**	−.08**	−.02*	NS	NS	NS
Foreign aid	.05**	.01**	.02*	NS	−.03*	NS	NS
Welfare	.04**	.06**	−.02*	NS	NS	NS	NS
Wealth	NS	NS	−.08**	NS	.03*	NS	NS
Average[a]	.06	.03	.02	.00	.00	.00	.00
	(.08)	(.04)	(.04)				

NOTE: All entries are *net* deviations from the mean on scales with a range of one. Positive entry indicates a desire for more government spending. Only deviations with $p < .05$ are entered. * = $p < .01$; ** = $p < .001$; NS = nonsignificant.

[a] Desire for less spending on defense is treated as a liberal response. Numbers in parentheses are average betas adjusted for independent variables and covariates.

ion, and since spending on drug use is not easily characterized in liberal or conservative terms, it seems safe to conclude that overall the young are generally more liberal in their views concerning government spending.

Urban residents demonstrate a pattern similar to, though more modest than, that of young people. The only real differences are that urban dwellers favor slightly more spending on social welfare and crime, and support slightly smaller increases in other policy areas (space, environment, education, foreign aid, and so on). With the exception of spending on drug abuse, however, these differences are of degree and not ideological direction: both urban dwellers and young people demonstrate the liberal pattern of greater spending in most policy areas except defense. Again, professionals show somewhat inconsistent patterns, though again the more liberal approach of greater spending in areas other than defense dominates. Professionals are, however, less likely to support spending on welfare or on equalizing wealth, suggesting that in the end, they know where their interests lie.

The mean deviations at the bottom of Table 6 summarize these trends nicely, showing that the liberal tendencies of yuppies are due largely to the main effects of age and residence. Over 80 percent of the distinctiveness of yuppies is accounted for by these two characteristics, and all of yuppies' liberal stands on government spending can be traced to the separate main effects of age, residence, and professional status. Again, the interaction of all three characteristics, the surest indication of a yuppie class effect, is unimportant.

All of the analyses presented thus far were based upon the entire 1972–1985 data file. While the social forces making yuppies a potentially distinctive political phenomenon have existed since at least the early 1970s, yuppies did not draw national attention until the 1980s. It is possible, therefore, that the yuppie thesis would fare better if only data from 1980 onwards were employed. Thus the statistical analyses were replicated once using only data from 1972 to 1979, and then using data from 1980 to 1985. The results strongly suggest that the interpretations presented above hold equally well for the 1970s and the 1980s. Of the 41 three-way interactions examined, only 2 (or 5 percent) were significant in the 1980s but not in the 1970s, a result easily attributable to chance. Five of the 123 two-way interactions (4 percent) were significant in the 1980s but not in the 1970s. Overall, 93 percent of the main, two-way, and three-way effects for the 1980–1985 data demonstrate the same levels of significance presented in Tables 4 through 6.

Conclusion

To summarize, despite the flood of media attention that has engulfed them, yuppies constitute a very small percentage of the population in general and of the baby-boom generation in particular. On strictly demographic grounds, then, there is reason to agree with the assessment that "the yuppies . . . don't deserve the attention they're getting" (American Demographics, 1985: 29). On the other hand, even though they are not especially numerous, yuppies do stand out from the great mass of Americans in their political attitudes and opinions: they are more liberal on a host of issues but less likely to identify with a political party, more likely to favor greater government spending on social and welfare programs, less likely to support defense spending, and considerably more politically tolerant than the general population. In light of their status relative to the population as a whole and to the new generation of Americans (as well as their perceived importance in the minds of journalists, politicians, and academics), they may constitute a more potent political force than their numbers alone would suggest.

The analysis of variance results indicates that yuppies' liberal tendencies are not due to some unique political profile that results from the combination of being young, urban, and professional, but instead are largely due to the separate, additive effects of these three characteristics. That is, yuppies are more liberal than the rest of the population because they are young, and young people are generally more liberal; because they are urban, and urbanites are generally more liberal; and because they are professional, and professionals are, on balance and in recent times, more liberal. If one is interested in identifying ideologically distinct portions of the population, one would be on much stronger ground examining the three larger categories defined by youth, urban residence, and professional status than their essentially coincidental intersection in yuppies, because it is largely the three main characteristics themselves that account for any political distinctiveness that exists.

The only evidence of any unique effects of the yuppies' defining characteristics is in the two-way interactions: on approximately a quarter of the issues we examined (30 of the possible 123), these interactions do affect the distinctiveness of yuppies beyond what would be expected from the main effects alone. These interactions, which are most common for young professionals, are about as likely to lead to a more conservative as to a more liberal yuppie political profile. Which two-way interactions are most common, as well as the directional pull of the effects, vary dramatically with the type of opinion, however. Effects on general issue stands (Table 4) are equally likely among the three types of interactions, but are exclusively liberal in their impact. For political tolerance (Table 5), 9 of the 10 two-way interactions are between age and professional status, and all 9 act to reduce tolerance. Two-way interactions are least common for opinions on government spending (Table 6), and are as likely to lead to increased conservatism as liberalism. However, to return to the truly important part of the story, any interaction effects—two-way or three-way—are almost invariably less important than the corresponding main effects. In fact, in only 2 of the 49 analyses of variance is the three-way interaction among youth, urban residence, and professional status statistically significant, let alone substantively important.

Thus, we have uncovered virtually no evidence that would support the yuppie class thesis, the notion that a synergism among the component elements of yuppiedom makes yuppies somehow *sui generis* politically. The generally liberal stands of yuppies are explainable without reference to any yuppie class or new class notion, and a very traditional group politics thesis performs quite nicely indeed. In certain instances we have also turned up evidence that points toward the applicability of the special effects thesis, for the conservative effects of professional status do occasionally come into conflict with the more liberal tendencies associated with youth and urban residence. The tendency of the two-way interaction effects to vary in their ideological impact depending upon the nature of the issue involved might also be interpreted as support for this thesis. But again, this is only a secondary pattern, greatly overshadowed by primary, main effects we have observed.

Of course, it is still possible that the yuppie thesis is of political import, but in ways that have not been tapped in this study. The concept of "a yuppie" may be less a matter of demographic characteristics than—to return to the third and fourth parts of Piesman and Hartley's definition—of a state of mind or a lifestyle. As we mentioned at the outset of this article, however, such an amorphous definition tends to confuse the consequences of the concept with the concept itself. For "yuppie" to be a useful construct in understanding contemporary politics, it must be defined with precision. And the research introduced here suggests that as the definition of yuppies becomes less amorphous, the political significance of yuppie status becomes more ephemeral.

References

American Demographics
1985 "The big chill (revisited), or whatever happened to the baby boom?" American Demographics 7 (September): 22–29.
Bell, Daniel
1973 The Coming of Post-Industrial Society. New York: Basic Books.
1979 "The new class: a muddled concept." Pp. 169–90 in B. Bruce-Briggs (ed.), The New Class? New Brunswick, NJ: Transaction Books.
Bennett, W. Lance
1980 Public Opinion in American Politics. New York: Harcourt Brace Jovanovich.
Brint, Steven
1984 " 'New-class' and cumulative trend explanations of the liberal political attitudes of professionals." American Journal of Sociology 90 (July): 30–71.
Bruce-Briggs, B. (ed.)
1979 The New Class? New Brunswick, NJ: Transaction Books.
Delli Carpini, Michael X.
1986 Stability and Change in American Politics: The Coming of Age of the Generation of the 1960s. New York: New York University Press.
Ehrenreich, Barbara, and John Ehrenreich
1977 The professional–managerial class. Radical America 7 (March-April): 7–31.
Gopoian, David, Derek Hackett, Daniel Parelman, and Leo Perrotta
1985 "The Democratic Party coalition in the eighties: a reassessment of Ladd's old class/new class explanation of party conflict." Paper, annual meeting of the Southern Political Science Association, Nashville.
Gouldner, Alvin W.
1979 The Future of Intellectuals and the Rise of the New Class. New York: Seabury.
Harrington, Michael
1979 "The new class and the left." Pp. 123–38 in B. Bruce-Briggs (ed.), The New Class? New Brunswick, NJ: Transaction Books.
Inglehart, Ronald
1977 The Silent Revolution: Changing Values and Political Styles among Western Publics. Princeton: Princeton University Press.
Kristol, Irving
1972 "About equality." Commentary 54 (November): 41–47.

Ladd, Everett C., Jr.
1978 "The new lines are drawn: class and ideology in America. Part I." Public Opinion 3: 48–53.
1979 "Pursuing the new class: social theory and survey data." Pp. 101–22 in B. Bruce-Briggs (ed.), The New Class? New Brunswick, NJ: Transaction Books.
Markus, Gregory, and M. Kent Jennings
1985 "Of yuppies, enddies, and the formation of political generations." Paper, annual meeting of the American Political Science Association, New Orleans.
Miller, Warren, and Teresa Levitin
1976 Leadership and Change. Cambridge, MA: Winthrop.
National Opinion Research Center
1972– General Social Surveys. Chicago: NORC, University of Chicago.
1985
Newsweek
1984 "The year of the yuppie." Newsweek (December): 14–20.
Piesman, Marissa, and Marilee Hartley
1984 The Yuppie Handbook: The State-of-the-Art Manual for Young Urban Professionals. New York: Long Shadow Books.
Podhoretz, Norman
1979 "The adversary culture and the new class." Pp. 19–32 in B. Bruce-Briggs (ed.), The New Class? New Brunswick, NJ: Transaction Books.
Price, Barbara A.
1984 "What the baby boom believes." American Demographics 6 (May): 30–33.
Reich, Charles
1970 The Greening of America. New York: Random House.
Stouffer, Samuel
1955 Communism, Conformity, and Civil Liberties. New York: Doubleday.
Wildavsky, Aaron
1979 "Using public funds to serve private interests: the politics of the new class." Pp. 147–54 in B. Bruce-Briggs (ed.), The New Class? New Brunswick, NJ: Transaction Books.

The Perceived Threat of Nuclear War, Salience, and Open Questions

HOWARD SCHUMAN, JACOB LUDWIG, AND
JON A. KROSNICK

FREEMAN DYSON begins his book *Weapons and Hope* (1984) with the statement: "I want to discuss the gravest problem facing mankind, the problem of nuclear weapons . . . ." Dyson's belief that the threat of nuclear war constitutes the most important problem facing Americans today has been shared by leading scientists from Albert Einstein to the 20 eminent signatories of a recent article on the consequences of a "nuclear winter" (Ehrlich et al. 1983), and has been echoed as well by a number of political and intellectual leaders. Their reasoning is straightforward: the United States and the Soviet Union have for many

---

Abstract The most serious problem facing the United States, according to many scientific and political leaders, is the threat of nuclear war. Yet the standard survey question on the most important problem facing the country has often shown little public concurrence with this assumption. Our article uses experimentation in national samples to test whether this difference can be traced to limitations in either the form or the wording of the standard question. The results indicate that there are some important systematic differences between open and closed versions of the question, and also differences that result from reference to the nation as distinct from the world, but neither type of difference accounts for the infrequent mention of nuclear war on the standard question. Instead, other evidence indicates that most Americans believe that nuclear war is not going to happen at all, or that if it does happen it will be too far in the distant future to be of pressing concern to them personally.

Howard Schuman is Director of the Survey Research Center and Professor of Sociology at the University of Michigan. Jacob Ludwig is Statistician and Methodologist at the Gallup Organization. Jon A. Krosnick is Assistant Professor of Psychology at the Ohio State University. An earlier version of this article was presented at the Annual Meeting of the American Association for Public Opinion Research, May 1984. The authors wish to thank Edward Blair for having raised a stimulating challenge to their analysis, and Lincoln Moses for providing statistical consultation that helped them deal with the challenge. The research on which the article is based was supported by a grant from the National Science Foundation (SES-8411371). The article was completed while the senior author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, with support from a National Science Foundation grant (#BNS-8011494).

Public Opinion Quarterly Vol. 50:519–536 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-519/$2.50

years and with ever-increasing certainty possessed sufficient nuclear weapons to destroy much of human civilization and perhaps extinguish the human race itself. In the face of such potential disaster, national issues, interests, and ideologies must be subordinated to the simple need for survival.

However, if we are to judge by surveys that allow spontaneous ranking of this issue relative to other issues, a strong concern about nuclear war has seldom been uppermost in the minds of the great majority of Americans. Surveys conducted during 1980 and 1981 by the University of Michigan Survey Research Center (Miller, 1981) and by the Gallup organization (Gallup, 1982) found fewer than 10 percent of the American adult population mentioning the threat of war of any kind when asked one or another variant of the basic question: "What do you think is the *most* important problem facing this country today?" During the 1950s and 1960s international issues did figure more prominently in answers to this standard index question, but for the most part these seemed to revolve around particular crises such as Vietnam, rather than around the more general issue of nuclear war (Smith, 1980). The two themes—nuclear war and more specific military engagements—are not easily separated, as we will later see, but the virtual disappearance of international categories from the Gallup results after the United States withdrew from Vietnam in the early 1970s shows how little the nuclear war threat itself was the enduring and overriding issue elicited by the most important problem question.

In this article we investigate two possible explanations for the relatively low spontaneous mention of the threat of nuclear war by Americans. One explanation implies methodological artifact: the standard open-ended most important problem question may itself be too limited, obtaining superficial answers that reflect what is momentarily salient to respondents about national events, rather than eliciting their more considered judgments. A finding by Schuman and Presser (1981) that a sudden cold spell in 1977 drastically affected responses to the most important problem question can be interpreted in just this way. The second explanation is quite opposite, however, for it accepts the survey data at face value and attempts to determine why the threat of nuclear war may not be regarded as especially important by most Americans, no matter how the question is asked.

We tested the two explanations by means of experimental comparisons of open and closed question forms at five time points over a 21-month period in 1982 and 1983, by follow-up probe questioning at two of the time points, by altering the wording and therefore the frame of reference of the questions, and by a review of some key indicators of the news environment for the same 18-month period. The article also explores more generally the possible distinction between *salience* and *importance* in survey questioning. In doing so, we add to ideas and evidence developed earlier by Schuman and Presser (1981) concerning the differences between open and closed questions in survey research.

Salience and Importance in the Study of Attitudes

The standard "most important problem" (hereafter, MIP) question that is used by many survey organizations to track changes in public perceptions of national issues requires a spontaneous reply on the part of the person answering, rather than a choice among offered alternatives. It is useful to think of the respondent's mental process in answering the question as a two-step sequence: a set of one or more issues must be brought to mind, then a choice made among them if there is more than one.

In the context of the survey interview, the first step in the process might seem to depend heavily on what is salient in the respondent's immediate environment, for example, what he or she heard most recently on television news, saw on a newspaper's front page, or experienced personally on the job, shopping in a store, or in some other recent life activity. It may be that inflation, unemployment, and other non-nuclear issues are frequently spotlighted in just these ways, but that the threat of nuclear war comes to mind less easily because much of the time it is not in the news a great deal—at least not in a concrete, easily visualized way—nor is it connected directly to daily life.[1]

This line of explanation is consistent with the dominant use of the term *salience* in cognitive psychology, where it is ordinarily employed to refer to conspicuous but not necessarily important aspects of the external world (for example, an odd hat) that catch a perceiver's immediate attention. Even where the internal world of memory is the reference, the stress is still on the fleeting, so that the phrase "seemingly trivial but highly salient information" (Taylor and Fiske, 1978:252) captures well the emphasis in much cognitive writing.

To be sure, salience has come to suggest matters of greater substance within survey research. Cannell and Kahn (1968) describe reporting accuracy as having "to do with the importance or salience of [an] event for the person reporting," thus using the two terms interchangeably. And Scott (1968), dealing with the kinds of attitude questions frequently used in surveys, suggests that salience is "perhaps indistinguishable from . . . the importance of the focal object to the person." Yet it may be that open survey questions, even questions that ask for the "most important problem," obtain mainly what is salient in the sense of conspicuousness, not what is deeply important to the individual.

---

[1] The effect of the mass media in influencing the definition of important issues for the public has been documented in a number of studies of agenda setting. See, for example, Iyengar et al. (1982), MacKuen (1981), McCombs (1981).

A closed question, on the other hand, by first offering a set of choices and then asking for a decision as to the most important, avoids the problem of differential salience among the issues presented. The respondent's sole task is to choose among the offered alternatives on the basis of considered judgment as to their relative importance.[2] If the reason that the threat of nuclear war is seldom given to the open MIP question is its lack of salience, then when added as a choice to a closed set of alternatives that are already given frequently to the open question, nuclear war should show a significantly larger (ideally, much larger) open-to-closed increment than the others.

We will first test this hypothesis directly, then consider the alternative line of explanation that takes seriously the apparent lack of public concern about nuclear war implied by responses to the open form of the MIP question.

Method and Preliminary Results

Open and closed forms of the MIP question were administered in split-ballot experiments to random halves of national samples in each of five monthly Survey Research Center (Survey of Consumer Attitudes) telephone surveys between April 1982 and December 1983. The exact wording of the questions and the coding categories used for the open form (and where relevant for the closed form) are presented in Table 1. The closed alternatives were developed in early 1982 by drawing on the largest categories reported for recently employed versions of the open question by the Gallup organization, with the restriction that no more than five closed alternatives could be used because prior experience suggested this to be the maximum number feasible in telephone surveys. In fact, after starting a priori with "the threat of nuclear war," only three other problems ("the high cost of living," "un-

---

[2] It might seem as though the order in which closed alternatives are read could influence responses, but Schuman and Presser (1981: 61–63) provide experimental evidence against this possibility with a very similar question. Furthermore, the same authors (1981: 99, 292–93) offer evidence against the easy assumption that small differences in phrasing closed alternatives necessarily affect distributions. Although the possibility of such effects remains in any particular instance and makes further empirical tests always desirable, we proceed here on the assumption that neither response order nor minor aspects of wording has a major impact on the univariate distribution to the closed question.

Table 1. Responses to Open and Closed Forms of the Most Important Problem Question[a]

	April 82		*June 82*		*October 82*		*July 83*		*December 83*	
Response	*Open*	*Closed*	*Open*	*Closed*	*Open*	*Closed*	*Open*	*Closed*	*Open*	*Closed*
High cost of living, inflation	16.1%	22.3%	16.1%	22.5%	5.8%	21.7%	5.3%	20.0%	2.1%	15.3%
Unemployment	28.0	42.6	28.2	39.3	42.9	45.5	41.5	45.3	22.3	28.0
Nuclear war, war	5.9	14.7	8.0	16.2	3.7	12.2	6.3	13.5	23.8	30.3
Budget cuts	1.6	7.6	2.9	10.5	1.0	3.2	—	6.5	0.7	10.3
Interest rates	1.1	0.5	3.4		1.6				0.4	0.4
Government deficit	1.6		0.6	1.0	1.6	3.7	1.0	2.4	2.1	4.2
General economics n.e.c.	9.1	0.5	9.2	1.6	11.0	1.1	6.8	0.6	7.1	1.5
Soviet aggression	1.1				0.5				0.7	
National defense	1.1	0.5	1.7	0.5	1.0	0.5	1.4		2.1	0.8
U.S. involvement in specific countries (e.g., Lebanon)	1.1				0.5		2.4		6.0	
Foreign affairs, n.e.c.	3.2		5.7	1.0	4.2	0.5	7.2	1.8	5.7	0.8
Crime	4.3	2.0	4.0		3.1	1.1	1.0	0.6	2.8	0.4
Food, energy shortages	1.1		1.1		1.0		1.4			
Moral, religious breakdown	5.4	2.5	5.7	3.1	6.8	2.1	7.7	2.9	7.4	0.8
Distrust of government	2.7	1.5	4.0	1.6	3.7	0.5	1.9	1.2	3.9	0.4
Other government defects	3.8	1.5	4.0	0.5	1.6	3.7	2.4	1.2	2.1	1.9
Other/more than one coded response	10.2	2.5	4.6	1.0	9.4	3.2	10.6	4.1	7.8	4.6
Don't know	2.7	1.0	0.6	1.0	0.5	1.1	2.9		2.8	0.4
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
N	186	197	174	189	191	189	207	170	282	261

[a] Open Question: "What do you think is the *most* important problem facing this country today?"

Closed Question: "Which of the following do you think is the *most* important problem facing this country today—the high cost of living, unemployment, the threat of nuclear war, or government budget cuts—*or* if you prefer, you may name a different problem as most important."

employment," and "budget cuts") were mentioned often enough at that point (February 1982) to seem to require inclusion as closed alternatives, with the first two the clearly dominant open responses at the time. ("Budget cuts" was far behind cost of living and unemployment, and just slightly ahead of responses under the general heading of "war.") We shall use the term *common categories* to refer to these four alternatives, since they were presented as choices on the closed form and also served as code categories for the open form. Other noncommon categories were developed as needed to accommodate other types of responses that did not fit the four common categories.

The closed version of the MIP question differed from the typical closed survey question by providing an explicit legitimation of "other" responses: it included after the four alternatives the statement "*or,* if you prefer, you may name a different problem as most important." This was done in order to avoid creating question form differences by fiat. Thus the total set of open and closed categories parallel each other to the extent that the two forms elicited the same types of responses. The legitimation of "other" responses to the closed question led to an average of 15 percent of the answers falling outside the four common categories with little variation (11–17 percent) across the five surveys. The 11–17 percent contrasts with figures of less than 2 percent in earlier experiments where such legitimation was not provided (Schuman and Presser, 1981).

Despite this amendment to the closed question, the open form produced many more noncommon category responses than did the closed, though most of these led to additional categories that were either very small or else were vaguely miscellaneous in nature (e.g., "general economic conditions, not elsewhere classified"). Furthermore, the overall proportion of open responses that fell into the four common categories considered as a totality is remarkably constant over the 21-month period: for the five surveys the percentage in the common categories on the open form varies within the narrow range of 49 to 55 percent, with no discernible trend up or down over time. In sum, even though proportions among the four common categories in Table 1 shift markedly over the 21 months, the open/closed comparisons deal, so far as we can tell, with much the same subpopulation (i.e., that part of the national population located within the common categories on each form) at all five time points.[3]

Main Results

OPEN/CLOSED DIFFERENCES IN MENTIONS OF NUCLEAR WAR

The hypothesis that the nuclear war threat is seldom mentioned to the open question because of its low salience cannot be adequately tested by simply comparing the open and closed percentages for this category in Table 1. Each of the four common categories shows an increase in percentage between open and closed forms, which will usually be the case as one moves from an open question to a closed question constraining respondents to a small number of specified categories. A more meaningful test of the hypothesis is to ask whether the nuclear war choice shows a greater open to closed form increment than do the other three categories.

---

[3] The one noncommon category of some frequency was "breakdown of morals and religion," but it also changes little in its proportion over the period we tracked, and we therefore assume that movement between it and the common categories was not an important influence on the latter during this time period.

To determine whether low salience does in fact reduce nuclear war responses on the open form, we use the chi-square test to compare the distributions for the two forms, with the null hypothesis being that salience is *not* a factor on the open form and therefore that the two distributions will not differ. Furthermore, in order to focus specifically on the nuclear threat alternative as against all others, the other three common categories are collapsed, creating a 2 by 2 chi-square test for each of the five years. When such a test is carried out, in three of the five surveys there is no significant difference between the two forms. Moreover, in one of the two months in which the forms do differ significantly ($p < .05$ in October 1982, and $p < .02$ in December 1983), the ratio of nuclear war responses to the other three collapsed categories is actually higher on the open form than it is on the closed form (December 1983). These results argue against the hypothesis that an artifactual limitation in the open question accounts for infrequent mention of the nuclear war threat on that form.

In addition, looked at in more absolute terms, in none of the first four surveys does the proportion of respondents choosing nuclear war as the most important problem on the closed form ever reach one out of five, whereas unemployment claims almost half the respondents in two of those months. This hardly suggests that respondents, when reminded of the nuclear war threat, gravitate toward it in great numbers. To be sure, in the fifth month the percentage choosing nuclear war on the closed form rises to 30 percent, but at that point the parallel open form common category shows almost as high a proportion. In sum, the low salience explanation for infrequent mention of nuclear war on the standard MIP question is not confirmed by the data in Table 1.

THE UNIMPORTANCE OF NUCLEAR WAR FOR THE AMERICAN PUBLIC

In four of the five closed question data sets shown in Table 1, the threat of nuclear war went unchosen by more than four-fifths of the population sampled. If this is not due to the form of the survey question, then it becomes plausible to conclude that much of the public does not share the belief of Dyson and others concerning the paramount importance of the nuclear threat. In an attempt to understand this relative lack of concern about nuclear war, we asked those respondents in our last two surveys who did *not* choose nuclear war on the

Table 2. Reasons for Not Choosing "Threat of Nuclear War" on Closed Questions[a]

Reasons	*Percent*
1. Denial that nuclear war is possible E.g., "I don't think there will ever be a nuclear war. It is no threat. One country will not destroy the other or take a chance."	20.9
2. Nuclear war not an immediate problem E.g., "Because the threat of nuclear war is sort of remote, but unemployment is a problem right now in this area."	25.0
3. Nuclear war is out of my control E.g., "Well, you've got to live while you're alive. It's something that we can't do anything about."	8.1
4. Combination of 1 and 2	7.8
5. Combination of 1 and 3	1.3
6. Combination of 2 and 3	5.3
7. Repeat of original response to closed question (E.g., "It's very difficult for young people to acquire homes and cars with inflation as it is.")	17.2
8. Other	10.9
9. D.K.	3.4
	100.0
(*N*)	(320)

[a] The question read: "We are also interested in *why* people think one problem is more important than another. For example, can you say why you think [R's answer to closed question] is more important than the threat of nuclear war?" Data are combined from July and December, 1983; the distributions of responses for the two months do not differ significantly.

closed form why they made a different choice.[4] A coding of their open answers to this follow-up question is provided in Table 2.

Although such after-the-fact explanatory answers must be treated with caution, they seem in this case to provide some genuine insight into why the threat of nuclear war is mentioned so infrequently in Table 1. If we put aside the proportion (17 percent) who simply reiterated that the problem they chose (unemployment or inflation in most cases) was a very important one, there were three major explanations, which were sometimes given in combination.

The most frequent explanation was that nuclear war is something to worry about for the distant future, but that unemployment/inflation/budget cuts is an important problem here and now. "I know nuclear war could happen, but it's not an everyday problem," said one respondent. The second most frequent explanation given was out-and-out denial that a nuclear war is possible: "Nobody is crazy enough to start a nuclear war." Finally, the third type of explanation frequent enough to call for a separate category was that the problem is one the respondent can do nothing about. This last explanation, which was the least frequent, is the only one where the threat of nuclear war might possibly be seen as judged truly most important by the respondent despite not being given as a spontaneous answer. The other two substantive categories really involve denial, the one in an absolute sense, the other by projecting the issue into a fairly distant future. In the latter case, the respondent may not be much more worried about nuclear war than he or she would be on reading about the likelihood of another ice age in some thousands of years. Even where the probability of future disaster is more imminent, as in the case of earthquakes for Californians, there is not much evidence that this creates *continuing* concern on the part of the local population.

---

[4] The same follow-up question was asked in both July and December, 1983. The proportions giving various answers did not differ significantly between the two months, despite the higher frequency of the nuclear war choice in the latter month, and we have therefore combined the two samples of follow-up responses in Table 3 in order to provide a more adequate number of cases for review. The exact wording of the follow-up question is given in the footnote to Table 2.

The explanations for unemployment or inflation that are merely repetitions of the original choice also suggest that the respondent does not treat the threat of nuclear war as a very serious possibility when confronted with it, though we should also note that such repetitions tend to be given disproportionately by those with lesser education. Considering other demographic correlates, the most striking is that for sex: men are nearly twice as likely as women to deny the possibility of a nuclear war ($p < .001$), and women are nearly twice as likely as men to say that a nuclear war is out of anyone's personal control ($p < .01$). Seeing nuclear war as a distant problem is not disproportionately characteristic of either sex, but it is significantly more frequent among younger than among older persons. (The same is true for more as against less educated persons, though age is the stronger correlate here.) None of these correlates modifies the basic outcome of the follow-up inquiry: respondents have little difficulty explaining why they do not consider the threat of nuclear war to be the most important problem facing the country today.

CHANGING THE FRAME OF REFERENCE: FROM NATION TO WORLD

We took one further step to make certain that the MIP question was not inadvertently missing potential nuclear war responses. Since the question refers to problems facing "this country today," it is conceivable that respondents are led to think primarily of domestic or purely national issues, whereas nuclear war may come to mind mainly within a larger international frame of reference.

To check this possibility, we carred out a split-ballot experiment in November 1985 in which the standard open MIP question was paired with a version identical in every respect except for substituting the phrase "the world" for "this country." With this new phrasing ("What is the most important problem facing the world today?"), respondents concerned about the threat of nuclear war are provided full opportunity—in fact, almost encouraged—to give such an answer.

The national telephone sample for this new experiment was small, but the results are unequivocally negative. Of the 56 people who received the standard "country" version, only 16 percent mentioned nuclear war; of the 46 people who received the "world" version, 19 percent mentioned nuclear war—a trivial and nonsignificant ($X^2 = 0.2$) increase. The new version of the question did produce more apparently world-related responses like "hunger," and it did not elicit typically domestic economic answers like "unemployment." Yet "nuclear war" did not increase appreciably as a result of the broader international frame of reference.

Overall, our experimental and related follow-up investigations show that it is incorrect to assume that the nuclear war threat has an underlying importance for Americans not revealed in their open responses. The main reason that nuclear war was not given as a response to the open question in much of 1982 and 1983 appears to be that it was not seen to be of paramount importance by respondents, not that it failed occur to them. Even when it was suggested as part of a set of closed alternatives, its selection did not increase greatly. And when asked why not, most respondents reported that it was either not a real problem or not a pressing one as compared with the other problems they saw. Finally, when encouraged to think of problems at the world level, there is still little or no increase in mentions of nuclear war.

THE SUDDEN EMERGENCE OF THE PERCEIVED NUCLEAR THREAT

Between July 1983 and December 1983, "the threat of nuclear war" category rose by 17 percent on the open form of the question and by 16 percent on the closed form (see Table 1). Why did this large increase occur? One of the best publicized news events about nuclear war in the last months of 1983 was the television showing of a film, *The Day After*, about the effects of such a war. The film received so much advance publicity that the Secretary of State appeared on television immediately after its showing in order the justify government policy on arms control. Moreover, an estimated 100 million people watched the program, one of the largest television audiences in the country's history (*The New York Times*, November 27, 1983, p. 11). It seems plausible that the film increased the salience of nuclear war and thus accounted for at least some of the increase in the category on both forms.

The television film was shown on November 20 and our December survey began just three days later. We asked respondents who received the open form of the MIP question in December if they had seen *The Day After*. Some 57 percent said they had seen all or part of it, 43 percent none of it. But there is little or no evidence of any effect from the film on mention of nuclear war: 26 percent of those who saw the film mentioned nuclear war on the open form, while 21 percent of those who had not seen it gave the same mention (p = n.s.). Moreover, if the data are separated into 10-day periods, the difference between viewers and nonviewers was smallest (a mere 1 percent) in the first 10-day period. These findings do not prove that the program had no effect on the public, but they do suggest that if there was an effect it was due to the general publicity surrounding the film, not the actual viewing of the film itself. There is one other possibility worth noting: *The Day After* dealt with the consequences of a nuclear war, not at all with its likelihood; perhaps a major television broadcast focusing on the probability of a nuclear war would have more impact on public beliefs of the kind tapped by the MIP question.

A series of real events may well have done much more than publicity about the film to raise public concern about nuclear war between July and December of 1983: news reports about European mass resistance to nuclear missiles; the shooting down of a South Korean airliner by the Soviet Union on September 1 with the loss of 269 passengers and crew; the bomb attack that killed 241 marines in Lebanon on October 23; the invasion of Granada on October 25; and finally the Soviet break-off of arms control talks in late November and early December. At the same time, diminished concern over unemployment and inflation left much of the public (and of the news media) freer to focus on the threat of war. Assuming that all these factors played some role in the dramatic change in responses to the MIP question, both forms of the question picked up the change to about the same degree in our December survey.

MEDIA REPORTS

In order to try to measure indicators of the link between actual events and respondent perceptions of problems, we analyzed a detailed listing of the topics covered by the evening national news broadcasts for the three major television networks (ABC, CBS, and NBC) for each of our survey months plus the just preceding month.[5] Each news story was counted if it dealt with inflation, unemployment, nuclear war (including international talks aimed at reducing nuclear weapons), or

---

[5] Based on the Index in each monthly volume of the *Television News Index and Abstracts*, published by the Vanderbilt Television News Archives, Vanderbilt University Library.

Table 3. Number of Evening News Reports Dealing with Four Topics[a]

	Mar.–Apr. 82	*May–June 82*	*Sep.–Oct. 82*	*June–July 83*	*Nov.–Dec. 83*
Inflation	12	12	13	11	11
Unemployment	14	18	62	22	14
Nuclear war	71	46	29	63	104
Budget	88	62	33	10	8

[a] Indicators counted in the Vanderbilt Television News Archive Indexes

Inflation:	Reports on the Consumer Price Index or the Wholesale Price Index
Unemployment:	Reports on unemployment or employment
Nuclear war:	Reports on nuclear war, atomic weapons, missiles, disarmament, antinuclear demonstrations
Budget:	Reports on federal budget

budget issues. There are important limitations to these data, since we did not have access to broadcast tone or detailed content (e.g., reports on arms control talks might have been optimistic or pessimistic), and in the case of the budget, reports did not necessarily imply budget *cuts*. Still, the data are of some value if these qualifications are borne in mind. The results are presented in Table 3.

Nuclear war and related issues were mentioned frequently in news reports in the first half of 1982, then dropped to a lower level, only to rise to their highest frequency at the end of 1983. The last time point fits our trend survey data for December 1983, but the early time point is higher than the questionnaire responses might lead one to expect. It seems likely that it was the *conjunction* in late 1983 of the more general reports on Soviet–American tension (e.g., the Korean jet liner) and the increase in specifically nuclear-related reports (break-off of disarmament talks) that accounted for the abrupt rise at the end of that year in mentions of the threat of nuclear war on both open and closed forms. More routine reports touching on nuclear arms in early 1982 apparently did not greatly affect public definitions of important problems. These had been primarily stories about ongoing arms control negotiations and about the nuclear freeze movement, but they were not accompanied by direct evidence of military action or of sharply increased Soviet–American tension.

The media results support our earlier conclusion that the enduring threat of nuclear war is not seen as paramount in importance by the public. It is only when the possibility of war is heightened by concrete events that the nuclear threat becomes important, and even then it may be involvement in a particular war (as in Vietnam or potentially in Central America), rather than the threat of a nuclear war, that usually becomes the focal issue.

Other Open–Closed Question Differences

Although our primary concern with open-closed form differences stemmed from our interest in the nuclear war response, it is useful to draw more broadly on the five survey experiments to determine whether question form had effects in ways not directly connected with our hypothesis about the nuclear response as such.

At the gross level of ranking of issues by the public, the answer is no. For four of the five time points, the rankings of the four common category issues are identical on the open and closed question forms, and for the remaining time point (July 1983) there is only a single minor reversal between ranks 2 and 3. Thus, if interest is only in discovering which issues are most frequently chosen as "most important," our evidence is that the two forms will ordinarily lead to the same conclusions, assuming, of course, that the closed form includes as alternatives the main issues that come to the fore on the open form.

However, closer examination of the open-closed comparisons provided by Table 1 leads to the following additional conclusion: the less frequently an issue is mentioned spontaneously to the open MIP question, the greater its increase in choice on the closed question. At the extreme, "budget cuts" is a response that is never given by more than 3 percent of the population to the open question, but it is chosen several times more often when offered as a closed alternative. That this example is not simply an accident due to peculiarities of the budget cut issue, to sampling error, or to the use of percentage scales is shown by a systematic analysis of all common categories with frequencies transformed into logits. "Budget cuts" represents only the extreme case of what is a quite general effect in these data across categories and across months for any given category.[6]

A plausible interpretation of this effect is that the closed question form makes all responses equally salient and that some people choose as most important an alternative such as "budget cuts" that would not otherwise have come to mind. Who are the people most likely to do this? It is useful to think of the closed common category samples as composed of two parts: those people who would have mentioned one

---

[6] Space constraints prevent inclusion of this analysis, but a copy can be obtained by writing to the first author. It should be noted that had we not had the budget-cut category and not looked systematically at these effects, the open to closed increase for the nuclear war response might well have been confounded with the more general open to closed effect, since the nuclear war open category was relatively small in the first four surveys. In that case, the original hypothesis about artifactual constraint on the nuclear war response in the open form would have been treated as confirmed. As it is, we can see that the open–closed difference is a general one, likely to apply to any reasonably important issue that is seldom mentioned spontaneously to the open MIP question.

Table 4. Unemployment and Inflation Responses and Rates by Month and Question Form

Responses and Rates	*April 1982*	*June 1982*	*October 1982*	*July 1983*	*December 1983*
Unemployment					
Open form	28	28	43	42	22
Closed form	43	39	46	45	28
Inflation					
Open form	16	16	6	5	2
Closed form	22	23	22	20	15
Rate of unemployment	9.0	9.4	10.2	10.0	8.4
Rate of inflation	7.6	6.6	5.9	3.4	3.9

NOTE: Response percentages are taken from Table 1, rounded off. The inflation rate is that for the second month preceding the month shown; the unemployment rate is that for the month preceding the month shown.

of the four common categories even on the open question, and those who would not have given a common category open response but who feel constrained on the closed form to choose among the common categories. It seems likely that it is mainly the latter set of people who migrate disproportionately to the low-frequency common categories. Such an interpretation will require panel data for more definite confirmation, but we can predict that almost any intrinsically important but seldom mentioned issue will show a relatively large jump in popularity when added to a limited set of MIP closed alternatives.[7]

There is one other slighter effect that may also characterize differences between open and closed MIP question forms, though its reliability is less certain. The open question seems to be somewhat more sensitive to changes in both external events and media attention to these events. This is best seen by tracing over time the trends by form for the common categories of unemployment and inflation. The percentages are reproduced in Table 4, together with U.S. government data on the unemployment and inflation rates for the same period.[8]

---

[7] Another process that would tend to increase the size of the smaller common categories on a closed form would be random choices by part of the sample. However, such "guessing" would be expected to occur particularly among less educated respondents, yet the small "budget cuts" category does not show a systematic relation to education on either question form. Thus there is no evidence that guessing plays a significant role in these closed question choices.

[8] The unemployment rate is the civilian unemployment rate, which is seasonally adjusted. The inflation rate was computed by dividing the CPI-U (Unadjusted Consumer Price Index) for each month by the CPI-U for the same month during the previous year. This figure is therefore also seasonally adjusted. Both rates were reported early in each month-long period in which our survey took place, though one (inflation) is regularly based on data from two months earlier and the other (unemployment) from one month earlier.

Mention of unemployment rises on the open form from 28 percent in April 1982 to 43 percent in October 1982, then drops to 22 percent in December, with the response curve following the curve of the unemployment rate fairly well. Moreover, the major increase in mention of unemployment occurs at the point that the unemployment rate was in the process of "breaking" the 10 percent threshold, a change that received a great deal of attention from the media (see Table 3). On the closed form, however, unemployment shows little or no rise in mentions between April 1982 and July 1983, then a noticeable decline of 20 points in December 1983. With the exception of the last drop, the closed form seems to track somewhat less well the changes in unemployment rates that were an important part of the period under study.

Mention of inflation decreases on the open form from a high of 16 percent in April 1982 to a low of 2 percent in December 1983. The inflation rate itself declined fairly steadily over the period of our study, though perhaps equally important was the fact that it had dropped substantially from the double-digit levels of 1979 and 1980. On the closed form, however, there is little decline at all in the mention of inflation through July 1983, despite the continued drop in the inflation rate, and even the apparent decline of 5 percentage points on the December closed form is small and nonsignificant.

However, none of the open–closed differences over time for unemployment and inflation can be shown to be statistically reliable. If responses are regressed on rates, although the unstandardized coefficients for open responses are larger than those for closed responses for both unemployment (12.2 vs. 8.6) and inflation (3.2 vs. 1.3), the differences do not reach significance ($p > .10$) with this small number of time points. Nor can the more important specific interactions of response by form by time be shown to be significant, such as the greater jump for the open than the closed unemployment category in October 1982. Thus the form difference just outlined, though theoretically plausible and seemingly visible to the eye, remains uncertain in reliability.

Conclusions

Two general hypotheses were investigated in order to account for the apparent lack of public concern over the threat of nuclear war, as indicated by infrequent mention of nuclear war to the standard open question about the most important problem (MIP) facing the country today. The first hypothesis was that the nuclear war issue was seldom mentioned because it is less salient on a day-to-day basis than other issues, despite the fact that at a deeper level the public considers it to be of paramount importance. We tested this hypothesis by including in a specially created closed form of the MIP question the alternative "threat of nuclear war," on the assumption that a substantial proportion of respondents who did not think of it spontaneously would nevertheless recognize its importance and select it from among the four closed alternatives. Results of five open–closed comparisons failed to confirm this hypothesis: the proportion of nuclear war responses rose on the closed form, but not generally more so than did other alternatives included on the same form. Furthermore, altering the frame of reference of the question to emphasize world rather than national problems increased mentions of nuclear war by only a trivial and nonsignificant amount.

A second hypothesis was that the failure of nuclear war responses to be given frequently to the MIP open question means essentially what it says: over much of the past four decades, most Americans have regarded the threat of nuclear war as less important than other major problems. The lack of a systematic open–closed difference for the nuclear war category is consistent with this hypothesis, and in addition further evidence is found in respondent explanations for why they did *not* choose the nuclear threat alternative on the closed question. Many people said they did not believe a nuclear war to be possible, and many others claimed that the threat is too far in the future to compete with what they saw as current pressing problems. Although one cannot rule out the possibility that such denial may be serving a protective psychic function (Lifton and Falk, 1982), at this point the burden of proof should be on those who offer such an explanation. The fact that nuclear war responses increased sharply in frequency on *both* question forms in late 1983 suggests that Americans are quite capable of giving the responses when events occur that seem to make the possibility of a major war more immediate.

A secondary purpose of our study was to continue the investigation of open/closed question differences begun by Schuman and Presser (1981). In the present comparison of open and closed versions of the MIP question at five points in time, we did not find gross differences in the ranking of problems between question forms. However, we did identify one form difference that seems likely to be quite general: the smaller a category is on the open form, the larger its increase on the closed form. A second possible form difference, which is less reliable in these data, is the somewhat greater responsiveness of the open question to changes in external events, such as a sharp rise in the unemployment rate or in the amount of media attention to such a rise.

However, the role of salience per se in influencing answers to the MIP and similar questions is severely constrained by question wording that calls for judgments within the framework of "importance" to the nation. Mere conspicuousness of an issue in news broadcasts or elsewhere is not likely to turn it into a frequently cited concern unless the public judges it to be of genuine significance. Thus, for all the attention Americans lavish on sports, not a single respondent in our five surveys mentioned an issue connected with sports as "most important," nor for that matter were there references to front-page crime stories or sensational scandals. Thus if there is something wrong with the ranking of nuclear war as less important than more transient domestic issues—including one as fleeting as difficulties due to an extreme cold spell—this has to do primarily with limitations in the public's judgment in dealing with an intangible future, not with the questions that assess that judgment. Observers who have witnessed the proliferation and spread of nuclear weapons over the past 40 years with little public protest should not find this conclusion difficult to comprehend.[9]

References

Cannell, Charles F., and Robert L. Kahn
1968 "Interviewing." In G. Lindzey and E. Aronson (eds.), The Handbook of Social Psychology. Vol. 2, 2nd ed. Reading, MA: Addison-Wesley Publishing Co.
Dyson, Freeman
1984 Weapons and Hope. New York: Harper & Row.
Ehrlich, Paul R., et al.
1983 "Long-term biological consequences of nuclear war." Science 222: 1293–1300.
Gallup Poll, The
1982 News release. February 7.
Iyengar, S., M. D. Peters, and D. R. Kinder
1982 "Experimental demonstrations of the not-so-minimal consequences of TV news programs." The American Political Science Review 76: 848–58.
Lifton, Robert Jay, and Richard Falk
1982 Indefensible Weapons. New York: Basic Books.
McCombs, M.
1981 "The agenda setting approach." In D. Nimmo and K. Sanders (eds.), Handbook of Political Communication. Beverly Hills: Sage Publications.
MacKuen, Michael
1981 "Social communication and the mass policy agenda." In M. MacKuen and S. Coombs (eds.) More Than News. Beverly Hills: Sage Publications.

---

[9] Of course, media or other attention may be necessary (even though not sufficient) to help the public conceptualize an issue, so that having an "unemployment rate" or an "inflation rate" helps focus public attention. But an even more important factor in shifts in response to the MIP question is its relative nature: one issue may gain or lose in apparent importance simply because another loses or gains, just as the size of a news story is determined by what other stories are available at the same time. Quite likely such relativity accounts in good part for the fact that war-related responses to the MIP question have continued at high levels since 1983, as indicated by Gallup Monthly Reports (e.g., 30 percent in January 1986). The absence of high unemployment and inflation over recent months has allowed issues of war (though not necessarily nuclear war) to move to center stage with relatively little competition.

Miller, Warren E., et al.
1981 American National Election Studies Data Sourcebook: 1950–1980. Cambridge, MA: Harvard University Press.

Schuman, Howard, and Stanley Presser
1981 Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context. New York: Academic Press.

Scott, William A.
1968 "Attitude measurement." In G. Lindzey and E. Aronson (eds.), The Handbook of Social Psychology, Vol. 2, 2nd ed. Reading, MA: Addison-Wesley Publishing Co.

Smith, Tom W.
1980 "America's most important problem—a trend analysis, 1946–1976." Public Opinion Quarterly 44:164–80.

Taylor, Shelley E., and Susan T. Fiske
1978 "Salience, attention, and attribution: top of the head phenomena." In L. Berkowitz (ed.), Advances in Experimental Social Psychology. Vol. 11. New York: Academic Press.

Ecological Concern, Attitudes, and Social Norms in Voting Behavior

JAMES D. GILL, LAWRENCE A. CROSBY,
AND JAMES R. TAYLOR

BEHAVIORAL scientists have shown considerable interest in public attitudes toward issues of environmental quality. This interest coincides with the continuing development of an environmental movement pressing for ecologically responsible production and consumption as the means for controlling environmental problems. The research on environmental attitudes has focused on identifying the social bases of "ecological concern," developing and improving measures of such concern, and investigating the effects of concern on environmental planning, public policy making, and citizen's behavior (see reviews by Van Liere and Dunlap, 1980, 1982). Unfortunately, reviews of the environmental literature have regularly omitted the stream of research conducted by consumer behavior scientists. This research parallels the work reviewed by Van Liere and Dunlap (1980, 1982), even to the extent that articles fall into the same three categories.

Abstract Using a theoretical framework not previously tested in environmental research, this study investigates the indirect relationship between ecological concern and voting behavior. Results indicate that the effects of ecological concern are mediated by attitudinal, normative, and behavioral intention variables. The results also provide a partial test of the validity of Ajzen and Fishbein's Theory of Reasoned Action. Key environmental, public policy, and communications implications are discussed.

James D. Gill is Assistant Professor of Marketing and Lawrence A. Crosby is Associate Professor of Marketing, in the College of Business, Arizona State University, Tempe, Arizona 85287. James R. Taylor is the S.S. Kresge Professor of Marketing, Graduate School of Business Administration, University of Michigan, Ann Arbor, Michigan 48109. The authors gratefully acknowledge financial support provided by the College of Business, Arizona State University and by S.S. Kresge Research Funds. The authors also wish to thank an anonymous reviewer for technical contributions made to the manuscript.

Public Opinion Quarterly Vol. 50:537–554 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-537/$2.50

First, there have been numerous attempts to determine the antecedents of ecological concern, especially as they relate to the bases of market segmentation (e.g., Crosby and Gill, 1981; Murphy, 1978; Brooker, 1976; Henion and Wilson, 1976; Webster, 1975; Kinnear et al., 1974). Second, efforts have been made to develop or improve ecological concern measures (e.g., Antil and Bennett, 1979; Henion, 1976; Kinnear and Taylor, 1973). Third, several assessments of the impact of ecological concern on consumption and voting behavior have been conducted (Crosby and Taylor, 1982, 1983; Crosby et al., 1981; Lepisto, 1979; Henion, 1976; Kinnear and Taylor, 1973; Herberger and Buchanan, 1971; Kassarjian, 1971). Generally, the research indicates that ecological concern is related to, but is not highly correlated with, consumption behavior. Although often a good explanatory variable of group (market segment) behavior, prediction at the individual level proves difficult. This conclusion parallels that of social and behavioral scientists who have expressed reservations about the value of attitude data in environmental research (e.g., Endo and Nielsen, 1974).

The weak relationship between ecological concern and behavior may mean that it is not an important explanatory variable. A more interesting possibility, however, is that ecological concern has an indirect rather than direct influence on behavior, with its influence being gradually diminished through successive levels of intervening variables. This view is consistent with Dunlap (1975) and Mazmanian and Sabatier (1981), who argue that a person's willingness to act upon his environmental attitude is subject to certain mediating influences (e.g., environmental issue salience, situational circumstances, individual differences). The researcher's theoretical perspective, of course, determines the hypothesized intervening variables and their relationships.

Van Liere and Dunlap (1982) state that "since progress toward the solution of environmental problems is likely to depend more on proenvironmental behaviors than proenvironmental attitudes, the reasons for the weak relationship between environmental attitudes and behaviors and the conditions under which it can be strengthened clearly deserve examination."

Research Purpose

The purpose of this study is to investigate the possible indirect relationship between ecological concern and environmental behavior. The objective is to identify a set of intervening variables which mediate the effects of concern on behavior. In assessing this relationship, a theoretical framework not previously applied in studies of ecological concern will be used. This perspective is the multiattribute or expectancy–value approach for modeling attitudinal inflluence.

Ajzen and Fishbein's Theory

Perhaps the most widely researched model in consumer behavior is based on Ajzen and Fishbein's Theory of Reasoned Action (Page, 1982). Introduced in 1967 (Fishbein, 1967), the theory has been developed, refined, tested (Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975), and applied in numerous contexts, including voting behavior, family planning, occupational choice, and brand choice (see reviews by Farley et al., 1981; Ajzen and Fishbein, 1973). The theory posits that people consider the implications of their actions and form intentions before engaging in a behavior. According to the theory, the determinants of behavioral intention are attitude toward the behavior and subjective norm (perceived normative expectations of various important others in the social environment). The determinant of attitude is cognitive structure, i.e., beliefs that the behavior leads to certain outcomes (b_i) and the evaluations of those outcomes (a_i). The determinant of subjective norm is normative structure, i.e., beliefs about the normative expectations of specific referents (nb_j) and the motivation to comply with those expectations (mc_j).

Ajzen and Fishbein (1980) relegate attitudes toward targets or objects (e.g., *the environment*, people, institutions) to the role of external or exogenous variables.[1] An integral part of the theory (i.e., the sufficiency assumption) is that external variable influences on behavior and/or behavioral intention are completely mediated by the set of endogenous variables representing the model's normative and attitudinal components (Loken, 1982).[2] In terms of the effects of ecological concern on environmental behavior, the implication is that this relationship will be indirect and possibly limited in magnitude. At the same time, ecological concern should have a direct and therefore much stronger influence on cognitive and normative structure.

Despite the logical appeal of the Theory of Reasoned Action, empirical research has provided only partial support for the theory's stated relationships (Loken 1982; Page, 1982). Inconsistent results have been found with respect to the mediating role of intentions and the validity of the sufficiency assumption. For example, Bagozzi (1981) found the attitudinal components of the model influence behavior only through their effect on intentions, which is consistent with the theory. In contrast, Bentler and Speckhart (1979) found behavior was directly rather than indirectly influenced by attitude toward the act. In a later study (1981), they found that attitude toward the act had both direct and indirect effects on behavior (i.e., behavioral intention partially mediated attitudinal influence). However, conclusions drawn from the Bentler and Speckhart studies require qualification, as the authors did not operationalize the attitude model according to recommended procedures and were focused on habitual or addictive rather than reasoned behavior.

---

[1] In this study an external (i.e., exogenous) variable is defined as one whose variability is assumed to be determined by causes outside the model under consideration.

[2] Endogenous variables are defined as those whose variation is explained by exogenous and other endogenous variables in the causal model under consideration.

Tests of the sufficiency assumption have also produced mixed results. Crosby and Muehling (1982), for example, found that external variables do have a significant direct effect on behavioral intention. In fact, one of the five external variables tested had a direct effect, yet no mediated effect on behavioral intention. Conversely, Loken (1982) found that external variables do not have a significant effect on behavior above that mediated by attitude toward the behavior and subjective norm. Although the focus of these studies differed (i.e., behavioral intention versus behavior), conclusions drawn from their results raise new questions about the validity of the sufficiency assumption.

Other questions have been raised concerning the possible interdependency of cognitive and normative structure variables (Miniard and Cohen, 1979, 1981). Fishbein and Ajzen have acknowledged the possibility of interdependence through the inferential formation of cognitive (1975:304) and/or normative beliefs (1975: 304, 306, 314). While in theory an inference could be made in either direction, a causal link from normative expectations to cognitive belief is most consistent with prevailing notions about the social influence process. Also, Ryan (1982) presents empirical evidence that cognitive structure is more a function of normative structure than vice versa.[3]

A test of the theory's stated and implied relationships, and its ability to explain the link between ecological concern and environmental behavior (see Figure 1), should contribute to a resolution of these construct and nomological validity issues. Further, such a test may improve understanding of how generalized environmental attitudes are translated into specific environmental behaviors. The hypotheses expressed below were formulated from specific relationships suggested

---

[3] Conclusions drawn from the Ryan (1982) study must be interpreted cautiously as the author's models were tested mainly on the basis of the chi-square, and no goodness-of-fit indices were computed. Further, a small sample was examined and no models successfully passed the overall chi-square goodness-of-fit test.

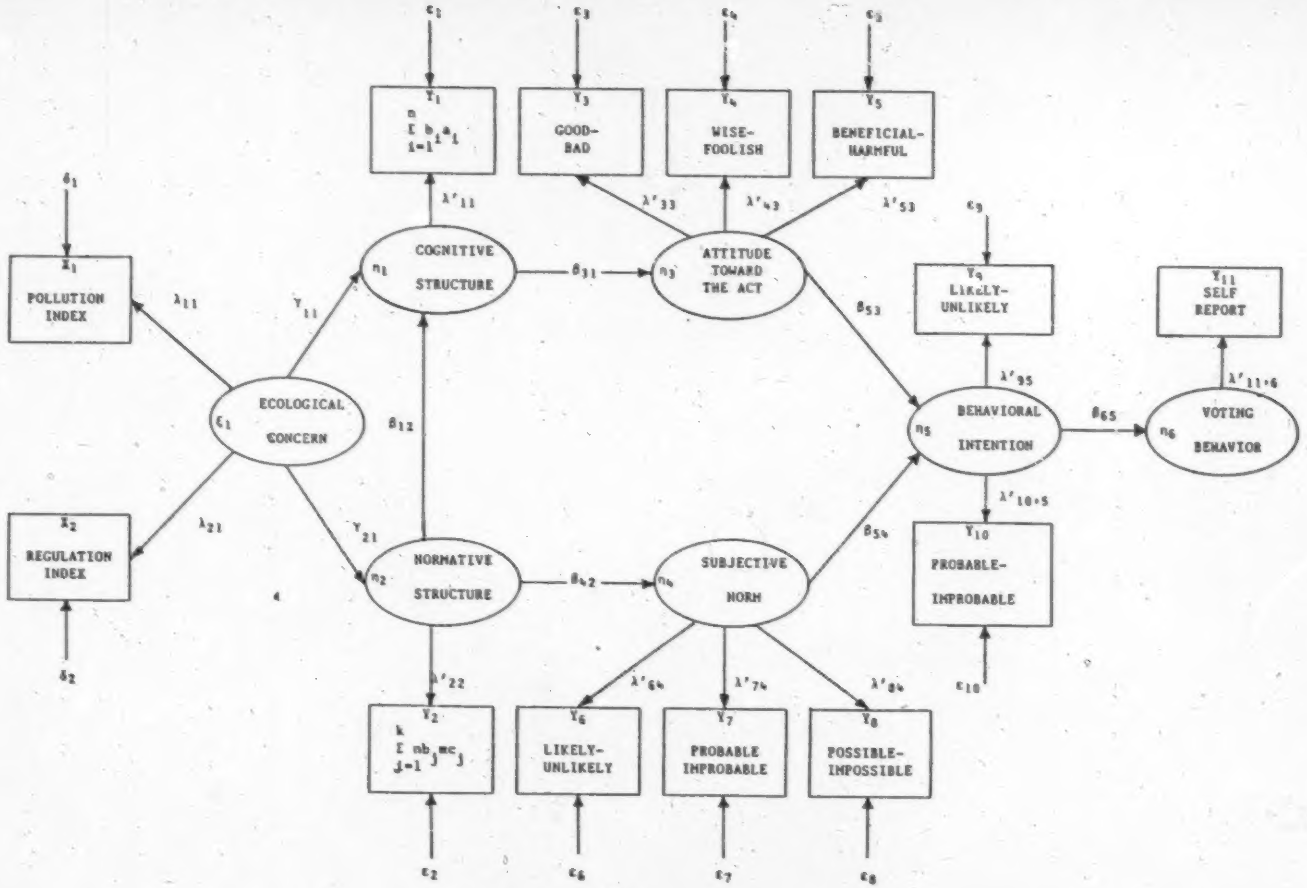

ε_1
ε_3
ε_4
ε_5
Y_1 $\sum_{i=1}^{n} b_i a_i$
Y_3 GOOD-BAD
Y_4 WISE-FOOLISH
Y_5 BENEFICIAL-HARMFUL
δ_1
X_1 POLLUTION INDEX
λ_{11}
γ_{11}
λ'_{11}
λ'_{33}
λ'_{43}
λ'_{53}
η_1 COGNITIVE STRUCTURE
β_{31}
η_3 ATTITUDE TOWARD THE ACT
β_{53}
ε_9
Y_9 LIKELY-UNLIKELY
Y_{11} SELF REPORT
λ'_{95}
$\lambda'_{11 \cdot 6}$
ξ_1 ECOLOGICAL CONCERN
β_{12}
η_5 BEHAVIORAL INTENTION
β_{65}
η_6 VOTING BEHAVIOR
$\lambda'_{10 \cdot 5}$
X_2 REGULATION INDEX
λ_{21}
γ_{21}
η_2 NORMATIVE STRUCTURE
β_{42}
η_4 SUBJECTIVE NORM
β_{54}
Y_{10} PROBABLE-IMPROBABLE
δ_2
λ'_{22}
λ'_{64}
λ'_{74}
λ'_{84}
ε_{10}
Y_2 $\sum_{j=1}^{k} nb_j mc_j$
Y_6 LIKELY-UNLIKELY
Y_7 PROBABLE IMPROBABLE
Y_8 POSSIBLE-IMPOSSIBLE
ε_2
ε_6
ε_7
ε_8

Figure 1. Specification of A Priori Causal Model

by the Theory of Reasoned Action and previous environmental research:[4]

H1: Relationships between cognitive structure, attitude toward the act, normative structure, subjective norm, behavioral intention, and behavior are as suggested by the Theory of Reasoned Action (see Figure 1).

H2A: Ecological concern has an indirect, but no direct relationship with environmental behavior.

H2B: Ecological concern has an indirect, but no direct relationship with behavioral intention.

Method

DATA

The environmental attitude–behavior relationship was tested in a public policy/marketplace research setting. The fate of container laws or "bottle bills" is an environmental issue of current public and research interest. Since 1972, laws requiring deposits on beverage and other recyclable containers have been passed in nine states through voter initiatives and legislative action.[5] Continuing congressional interest in enacting federal legislation requiring container deposits has placed this issue squarely on the public agenda. The data for the analysis were from a study of voter/consumer behavior on bottle bills appearing on the November 2, 1982 ballots in California and Colorado. Although there were minor differences in the language and legislative effect of the two bills, they were essentially similar propositions.

A quota sample of households in these states was obtained from a national consumer panel. Households were selected for inclusion by matching four household characteristics with known census data for each state. The characteristics were population density, age of head, household income, and household size. A random selection procedure was used to select an eligible adult voter within each household. The questionnaire, preceded by a prealert postcard, was mailed 14 days before the election to 500 target individuals in each state. A standard inducement was offered to panel members. A total of 662 completed surveys were returned (a 66.2 percent response rate). In the week after the election follow-up telephone interviews were completed with 447 mail survey respondents and 100 nonrespondents to determine actual voting behavior (self-reported).

---

[4] Ajzen and Fishbein (1980) have suggested that external variables may moderate the relative influence of Aact and SN on BI. In statistical terms, this implies that the external variables interact with the endogenous variables to affect behavioral intention indirectly. However, no theoretical reason could be identified to suggest that ecological concern would have a moderating effect on behavioral intention or behavior.

[5] States with container deposit laws include Connecticut, Delaware, Iowa, Maine, Massachusetts, Michigan, New York, Oregon, and Vermont.

The sample was somewhat overrepresentative of adults registered to vote (79.7 percent). In 1982, the estimated percentage of registered adults was 68.3 percent in Colorado and 63.7 percent in California. Similarly, turnout of registered voters within the sample (94.2 percent) was higher than the estimated voter turnout in Colorado and California of approximately 70 percent.[6] A difference was also noted between the California vote projection of 51.6 percent yes, 48.4 percent no (based on 246 telephone respondents' self-reports) and the actual vote outcome of 44.1 percent yes, 55.9 percent no ($\chi^2_1 = 5.97; p < .05$). However, the projected vote outcome for the Colorado sample (based on 227 telephone respondents' self-reports), 27.7 percent yes, 72.3 percent no, does compare favorably to the actual vote outcome of 25.5 percent yes, 74.5 percent no ($\chi^2_1 = 0.60; p <$ ns).

MEASURES

The items of measurement included components of the Fishbein attitude model, attitude statements pertaining to ecological concern, involvement, lifestyle, and values orientation. Only the first two categories of information are relevant to this analysis.

In an effort to improve the measurement of ecological concern, researchers have turned to the use of multiple-item rather than single-item indicators and increasingly relied on standardized measures. However, as noted by Van Liere and Dunlap (1982) existing measures of ecological concern differ in terms of (1) the extent to which they incorporate different environmental issues, and (2) assumptions of what constitutes an "expression" of concern. Not surprisingly, early tests of the degree of intercorrelation among standardized measures produced mixed results.

More recently, Van Liere and Dunlap (1982) tested the intercorrelation among six different ecological concern scales. Three of the scales dealt separately with issues central to the environmental controversy (e.g., pollution, natural resources, and overpopulation). Additional indices were constructed to reflect different "expressions" of ecological concern (e.g., favor government environmental regulation, favor government environmental spending, engage in environmental protection behaviors). They found that the overpopulation index and environmental behavior scale were only weakly linked to the other four indices. Conversely, the pollution, natural resources, government regulation, and government spending scales were highly correlated. Van Liere and Dunlap (1982) concluded that "environmental concern may be a fairly broad concept, but one that is best represented by concern about pollution and natural resources."

---

[6] The noted differences may reflect nonresponse bias, self-report inaccuracies, or possible imperfections in the consumer panel (e.g., underrepresentation of highly mobile people). Although the authors do not believe these differences invalidate the investigation of hypothesized effects we cannot rule out selection artifact bias.

This study's approach to measuring ecological concern was patterned after Van Liere and Dunlap (1982). Since the intent was to measure concern as an attitude, all 10 items composing the Maloney et al. (1975) affect (A–) scale were included in the battery. Also included were the five attitude (as opposed to behavior) items from Kinnear and Taylor's (1973) Ecology Scale and the six attitude items from Antil and Bennett's (1979) Socially Responsible Consumption Behavior Scale. Each item had a seven-category response format ranging from "strongly disagree" to "strongly agree." Responses were coded from 1 to 7 (reverse coded where item wording required) so that higher scores indicated more concern. The 21 items in the item pool were then sorted on an a priori basis into two categories: those dealing with pollution and natural resource problems and those focusing on regulation and spending approaches to combating those problems (see Appendix). Separate scales formed by summing responses to the attitude items were found to have acceptable levels of reliability ($\alpha_1 = .86$ and $\alpha_2 = .86$, respectively).

Operationalization of the Ajzen and Fishbein model required the identification of salient outcomes and referents prior to the construction of endogenous variable measures. A free-elicitation technique common to this type of research (Ryan and Etzel, 1976) was employed in a pretest of 73 respondents. A content analysis of elicited responses identified 21 separate outcomes and 14 distinct referents (see Appendix) about which cognitive and normative measures were developed.

The measurement approach and format suggested by Ajzen and Fishbein (1980) was applied in assessing the endogenous variables specified by the Theory of Reasoned Action. However, multiple items were used to measure each model component so as to avoid bias from adjective specificity (Ryan, 1982). The adjective pairs that anchored the ends of the scales were taken from previous research. For example, the components of cognitive and normative structure were measured as follows:

(b_5) My voting "yes" on the beverage container law would reduce the amount of solid waste that has to be burned or buried.

likely ___:___:___:___:___:___:___ unlikely

probable ___:___:___:___:___:___:___ improbable

(a_5) Reducing the amount of solid waste that has to be burned or buried is . . .
good ___:___:___:___:___:___:___ bad
wise ___:___:___:___:___:___:___ foolish

(nb_{10}) Most other consumers think I should vote "yes" on the beverage container law.
likely ___:___:___:___:___:___:___ unlikely
probable ___:___:___:___:___:___:___ improbable

(mc_{10}) Generally speaking, I want to do what other consumers think I should do.
likely ___:___:___:___:___:___:___ unlikely
probable ___:___:___:___:___:___:___ improbable

Bipolar (+3 to −3) coding was used for all scales except for the motivation to comply items (scored 7 to 1). Ajzen and Fishbein (1980) have suggested this mc_j scoring method in response to criticism that such items are not viewed by respondents as bipolar (for details, see Miniard and Cohen, 1981; Ajzen and Fishbein, 1980; Fishbein, 1976). The index of cognitive structure was formed by averaging the scales representing each component, multiplying corresponding beliefs and outcome evaluations, then summing across outcomes ($\Sigma b_i a_i$). The index of normative structure was computed in an analogous fashion. The adjective pairs used to operationalize attitude toward the act, subjective norm, and behavioral intention appear as indicators (the boxes) in Figure 1.

Analysis and Results

Structural equation analysis (Jöreskog and Sörbom, 1981), as suggested by Andrews (1984), was performed to test the model and identify direct and indirect relationships between the exogenous and endogenous variables. The analysis was conducted using the LISREL VI program which provides a simultaneous test of the structural model (having the form: $\eta = \beta\eta + \Gamma\xi + \zeta$) and the measurement models for the y's and the x's ($y = \Lambda_y\eta + \epsilon$, $x = \Lambda_x\xi + \delta$). The objective of the analysis is to find a set of values for the model parameters which minimizes the difference between the actual and implied correlations among observed variables.[7] The program employs a maximum likelihood procedure to estimate all nonzero, nonfixed parameters. Paths without arrows in Figure 1 were assumed to be zero. A multigroup analysis, treating California (CA) and Colorado (CO) as separate groups, was performed to test and cross-validate the structural and measurement models.

---

[7] Separate analyses were performed using the correlation and covariance matrices as input. While use of the covariance matrix is generally preferred, problems were encountered in terms of not-positive-definite error matrices. These problems were not observed when the correlation matrix was used as input. The difficulty may stem from the (unavoidable) dichotomous coding of voting behavior. Also, the use of interval rather than ratio scale variables is technically a violation of the assumptions of LISREL analysis. Since certain latent variables will not be normally distributed on an a priori basis, the analysis might have been adversely affected.

To enable the program to run, two initial changes in the a priori model were required. First, three observed variables having extremely high correlations with other indicators of the same construct were removed from the analysis (Y_3, Y_6, Y_9). Second, to facilitate model identification, the error on voting behavior was fixed at zero ($\theta_{11,11} = .00$).

In the initial test of the a priori model each free parameter was specified to be equal over groups, i.e., to have the same value for California and Colorado. This, a most restrictive test of the model, yielded a fair degree of model fit (goodness of fit: CA = .892, CO = .915), but the model must be rejected on the basis of the chi-square goodness-of-fit test ($\chi^2_{84} = 202.53$, $p < .000$). Although the use of χ^2 as an absolute index of fit is open to question, and there is some agreement that it should be used as a guide rather than as a rule (see review by Bentler, 1980), an effort was made to improve model fit.

A second test of the a priori model imposed a less restrictive requirement that the same pattern of free and fixed elements would exist for both California and Colorado, but not necessarily the same parameter values. This test of the model resulted in an improvement of model fit ($\Delta\chi^2_{26} = 85.94$, $p < .001$; goodness of fit: CA = .918, CO = .965), but the model remained inadequate ($\chi^2_{58} = 116.59$, $p < .000$).

Next, a path linking ecological concern to behavior was included (γ_{61}). However, the addition of this path did not result in an improvement in model fit ($\Delta\chi^2_2 = 3.89$, $p <$ ns), and H2A was accepted. Subsequently, and instead, a path linking ecological concern to behavioral intention (γ_{51}) was included in the model. The result of this change was an improvement in model fit ($\Delta\chi^2_2 = 9.56$, $p < .01$; goodness of fit: CA = .924, CO = .967), and H2B was rejected. Despite the improvement in model fit, however, the model remained inadequate according to the traditional chi-square significance criterion ($\chi^2_{56} = 106.93$, $p < .000$). However, for large samples, some suggest that a 2:1 χ^2/d.f. ratio might be considered reasonable (Wheaton et al., 1977). With a combined sample of 379 complete data cases, the obtained 1.91 ratio seems to suggest the model is adequate or at least nearly so. At this stage of testing, none of the normalized residuals for Colorado exceeded 1.8 and 84 percent were less than one. For California, the largest normalized residual was 1.4, and 95 percent were less than one.

In a final effort to improve fit, some correlated measurement error among selected indicators ($\theta^{\epsilon}_{4,8}$, $\theta^{\epsilon}_{4,10}$) was allowed in the California

Table 1. Standardized Parameter Estimates

Structural Model					Measurement Model				
	California		Colorado			California		Colorado	
Parameter	EST	(T-Value)	EST	(T-Value)	Parameter	EST	(T-Value)	EST	(T-Value)
γ_{11}	.111	(1.90)	.119	(2.21)	λ_{11}	1.000	(*)	1.000	(*)
γ_{21}	.280	(3.45)	.410	(4.98)	λ_{21}	.756	(9.11)	.933	(11.71)
γ_{51}	.126	(3.05)	.062	(1.52)	λ'_{11}	.942	(*)	.942	(*)
β_{12}	.792	(11.32)	.843	(12.60)	λ'_{22}	.879	(*)	.879	(*)
β_{31}	.853	(15.58)	.943	(17.74)	λ'_{33}	**	(**)	**	(**)
β_{42}	.874	(13.06)	.897	(14.93)	λ'_{43}	.987	(*)	.987	(*)
β_{53}	.652	(12.80)	.880	(15.19)	λ'_{53}	.965	(30.74)	.963	(37.77)
β_{54}	.246	(5.07)	.019	(.34)	λ'_{64}	**	(**)	**	(**)
β_{65}	.924	(15.89)	.908	(25.83)	λ'_{74}	.978	(*)	.978	(*)
ψ_{11}	.348	(5.42)	.150	(3.61)	λ'_{84}	.973	(36.57)	.952	(24.95)
ψ_{22}	.910	(7.06)	.844	(6.92)	λ'_{95}	**	(**)	**	(**)
ψ_{33}	.227	(5.31)	.166	(4.81)	$\lambda'_{10,5}$	.984	(*)	.984	(*)
ψ_{44}	.250	(5.05)	.179	(4.34)	$\lambda'_{11,6}$	1.000	(*)	1.000	(*)
ψ_{55}	.224	(5.97)	.158	(6.13)	θ^{δ}_{11}	.111	(1.19)	.232	(2.73)
ψ_{66}	.330	(8.10)	.161	(7.22)	θ^{δ}_{22}	.480	(6.43)	.268	(3.38)
R^2_1	.662		.844		θ^{ϵ}_{11}	.095	(2.35)	.173	(5.26)
R^2_2	.079		.166		θ^{ϵ}_{22}	.241	(5.10)	.238	(6.01)
R^2_3	.767		.838		θ^{ϵ}_{33}	**	(**)	**	(**)
R^2_4	.751		.820		θ^{ϵ}_{44}	.050	(2.96)	.031	(2.97)
R^2_5	.769		.847		θ^{ϵ}_{55}	.097	(5.39)	.076	(6.13)
R^2_6	.667		.841		θ^{ϵ}_{66}	**	(**)	**	(**)
					θ^{ϵ}_{77}	.044	(2.68)	.070	(3.04)
					θ^{ϵ}_{88}	.049	(2.98)	.118	(4.92)
					θ^{ϵ}_{99}	**	(**)	**	(**)
					$\theta^{\epsilon}_{10,10}$	.074	(2.24)	.035	(1.99)
					$\theta^{\epsilon}_{11,11}$	.000	(*)	.000	(*)
					$\theta^{\epsilon}_{4,8}$	−.029	(4.15)	.000	(*)
					$\theta^{\epsilon}_{4,10}$	.039	(2.79)	.000	(*)

* Constrained or fixed parameter, not estimated.
** Parameter involves observed variable deleted from analysis.

group, without theoretical justification. This relaxation of the model had the desired effect of sufficiently improving the model's fit ($\Delta\chi^2_2 = 30.87$, $p < .001$), such that it accounted quite well for the observed correlations ($\chi^2_{54} = 76.06$, $p < .026$). Although the χ^2 was still rejected at the .05 level, other indicators of model fit (goodness of fit: CA = .956, CO = .967) suggest the model is tenable, and that freeing additional parameters would essentially involve overfitting. The parameter estimates for this final model are reported in Table 1. Particularly noteworthy, is the substantial amount of variance explained in the endogeneous variables except for normative structure. The explained variance is typically in excess of 80 percent for Colorado and 70 percent for California, with behavior being one of the variables accurately predicted.

As stated in H1, the endogenous variables were related in the manner specified by the Theory of Reasoned Action. That is, cognitive structure was linked to attitude toward the act (β_{31} = .853 for CA; = .943 for CO) which, in turn, influenced behavioral intention (β_{53} = .652 for CA; = .880 for CO). Normative structure was linked to subjective norm (β_{42} = .874 for CA; .897 for CO), which had a significant relationship with behavioral intention in California (β_{54} = .246) but not Colorado (= .019). A strong relationship between normative and cognitive structure was found to exist (β_{12} = .792 for CA; .843 for CO), suggesting that behavioral beliefs may be influenced by inference(s) drawn from the content of normative structure.

Environmental voting behavior was influenced by behavioral intention (β_{65} = .824 for CA; = .908 for CO), but was not directly linked to ecological concern. This result lends support to the sufficiency assumption. Conversely, ecological concern had a direct influence on behavioral intention, significant in the case of California (γ_{51} = .126) but not Colorado (= .062), which is inconsistent with the sufficiency assumption. It should be noted, however, that the relative strength of this association was slight.

Examination of the path coefficients connecting ecological concern to cognitive and normative structure reveals generally weak relationships in comparison to the relationships between the model's endogenous variables. The link from concern to cognitive structure was barely significant in Colorado (γ_{11} = .119) and nonsignificant in California (γ_{11} = .111). Normative structure was found to be somewhat more dependent on ecological concern in both states, although the coefficients were still modest (γ_{21} = .280 in CA; = .410 in CO).[8]

In an attempt to isolate the specific aspects of cognitive structure contributing to voter preferences, a stepwise discriminant analysis (Wilks) of yes and no voters was performed on the 21 belief times outcome evaluation products. The results indicate that 6 $b_i a_i$ products significantly contributed to the discriminant equation ($p < .05$). These outcomes were "reduce litter," "improve public health and safety," "block better solutions to the problem," "increase beverage costs," "give people, especially children, an incentive to pick up discarded containers," and "create more government regulation."

Normative considerations were also found to contribute significantly to the formation of intention, both directly and indirectly. Therefore, the 13 normative belief times motivation to comply products were entered into a stepwise discriminant analysis (Wilks) of yes and no voter groups to identify salient referents. Five contributed to the discriminant equation ($p < .05$). These referents were "members of my family," "environmentalists," "close friends," "journalists," and "other consumers."

---

[8] The nonsignificant t-value for θ_{11} is not statistically different from zero. Since it is incongruous to expect zero measurement error in a large survey, the possibility of specification error must be acknowledged.

Discussion

The analysis provided evidence to support or, more properly "failed to falsify," H1 (model validity) and H2A (sufficiency assumption–behavior). Evidence was also provided leading to the rejection of H2B (sufficiency assumption–intention). Collectively, these results provide an explanation of the relationship between environmental attitudes and behavior, so often debated in the literature. The supported hypotheses confirm the fundamental role of ecological concern as a background variable which indirectly affects behavior.

It can be concluded that the influence of ecological concern, measured as a generalized or global attitude, is mediated by more specific attitudinal, normative, and behavioral intention variables. This outcome should be of interest to researchers who have questioned the use of attitude data in environmental policy research. It appears that the effects of ecological concern have been misinterpreted and improperly specified in previous research. By incorporating intervening attitudinal and normative variables research perspectives are likely to become more focused, increasingly powerful, and conceptually enhanced.

At the same time, the results raise doubt as to whether a continued focus on generalized environmental attitudes is warranted. While these measures may have some value in group level analysis, in this study their performance in individual-level prediction was found to be poor. In contrast, more domain-specific measures such as those incorporated in the Fishbein model are able to account for the majority of the variance in behavior. By operationalizing attitude and behavior at the same level of specificity (in terms of action, time, target, and context), prediction is greatly enhanced without diminishing the importance of ecological concern. In fact, respondents' evaluations of outcomes (a_i) made salient by an environmental issue, may provide a more accurate indication of their ecological concern than scores on a very abstract environmental attitude index. By considering their evaluations across a set of interrelated and sometimes conflicting outcomes, a more realistic and multidimensional assessment of their ecological concern is obtained. Furthermore, this approach places greater emphasis on how information about an environmental decision is channeled by a variety of ecological and domain-specific values.

For those concerned with solving environmental problems, the theory identifies the immediate determinants of environmental behavior that can be used in developing strategies or tactics designed to increase a person's *willingness to act on his ecological concern*. For example, since attitude toward the action is determined by cognitive structure, behavior could be influenced by changing the person's behavioral beliefs (b_i), adding behavioral beliefs to the person's cognitive structure, altering the person's evaluation of outcomes (a_i), or impacting the relative salience the person associates with outcomes. Because subjective norm is determined by the person's normative beliefs and motivation to comply with salient referents, behavior could also be influenced by affecting perceptions of referent expectations (nb_j), identifying new referents, influencing the relative salience of referents, or affecting the person's desire to comply with social influence (mc_j).

Regarding tests of the causal relations among model components, the findings (e.g., support of H1) were generally consistent with Ajzen and Fishbein's writings. Although in one state (CO), subjective norm was found to have a nonsignificant impact on intention, this can be explained by situational factors influencing the relative importance of personal versus normative considerations. One such factor might be the nearness of the election day and the opportunity for considerable prior deliberation. The Theory of Reasoned Action acknowledges that these situational influences exist.

The path linking normative structure with cognitive structure may be seen as raising questions regarding the conceptual and/or operational separation between the attitudinal and normative influences. However, before concluding that the constructs are invalid, it should be noted that the presence of such a relationship, or the lack thereof, may also be dependent on the situation (Fishbein and Ajzen, 1981; Miniard and Cohen, 1981, 1979). Moreover, while normative and cognitive structure variables were linked, subjective norm was directly influenced only by normative structure, and attitude toward the act only by cognitive structure. This result supports the view that there is utility in separating attitudinal and normative variables, as suggested by Azjen and Fishbein (1980), despite the fact that they are correlated.

Tests of the sufficiency assumption produced mixed results. That is, the effect of ecological concern on intention was only partially mediated by the attitudinal and normative variables (rejection of H2B). However, the impact of this external variable on behavior was entirely indirect (acceptance of H2A). Interestingly, these results are very consistent with the contrary outcomes reported in earlier studies. For example, Crosby and Muehling (1982) found external variables had direct and indirect effects on intentions; however, they had no measures of actual behavior. Conversely, Loken (1982) found the influence of external variables on behavior to be completely mediated; but she did not report any test of effects on behavioral intention. The sufficiency assumption results reported in this study should help clarify previous findings; however, the reported outcome does not provide any satisfactory resolution of the issue.

Appendix

POLLUTION AND NATURAL RESOURCE ITEMS[9]

I feel people worry too much about pesticides on food products.

It frightens me to think that much of the food I eat is contaminated with pesticides.

I feel fairly indifferent to the statement—"The world will be dead in 40 years if we don't remake the environment."

I become incensed when I think about the harm being done to plant and animal life by pollution.

I'm usually not bothered by so-called "noise pollution."

I get depressed on smoggy days.

When I think of the ways industries are polluting, I get frustrated and angry.

The whole pollution issue has never upset me too much since I feel it's somewhat overrated.

I rarely ever worry about the effects of smog on myself or my family.

It is futile for the individual consumer to try to do anything about pollution.

I think that a person should urge their friends not to use products that pollute.

Every person should stop increasing their consumption of products so that our resources will last longer.

The benefits of modern consumer products are more important than the pollution which results from their production and use.

All consumers should be interested in the pollution aspects of products they purchase.

I am personally interested in the pollution aspects of products I buy.

Natural resources must be preserved even if people must do without some products.

REGULATION AND SPENDING ITEMS[9]

It genuinely infuriates me to think that the government doesn't do more to help control pollution of the environment.

The government should force all products off the market that pollute.

I don't think we're doing enough to encourage manufacturers to use recyclable packages.

Consumers should be made to pay higher prices for products which pollute the environment.

Commercial advertising should be forced to mention the ecological disadvantages of products.

OUTCOMES

1. Would reduce litter.
2. Would improve public health and safety.
3. Would block better solutions to problem.
4. Would increase beverage costs.
5. Would give people, especially children, incentive to pick up discarded containers.

[9] All items were answered on 7-point Likert scales labeled strongly agree, strongly disagree.

6. Would create more government regulation.
7. Would reduce solid waste that has to be burned or buried.
8. Would hurt the state's economy.
9. Would restrict individual freedom.
10. Would make it difficult to obtain the deposit refund.
11. Would penalize everyone for the mistakes of a few.
12. Would give litterbugs an incentive not to litter.
13. Would make those who litter pay for their own carelessness.
14. Would conserve energy and scarce resources.
15. Would give the average person an incentive not to litter.
16. Would add to my inconvenience in terms of storing and transporting empty containers.
17. Would protect the environment.
18. Would eliminate jobs in the container manufacturing industry.
19. Would reduce the assortment of brands and types of containers.
20. Would create jobs to handle recycling.
21. Would lower litter collection costs now paid for in taxes.

REFERENTS

1. Most members of my family.
2. The environmentalists.
3. My close friends.
4. Most journalists.
5. Most other consumers.
6. Most government officials.
7. The bottlers.
8. Most outdoor recreationists.
9. The liberals.
10. My co-workers.
11. Recycling center employees.
12. The conservatives.
13. The retailers.

References

Ajzen, I., and M. Fishbein
1973 "Attitudinal and normative variables as predictors of specific behaviors." Journal of Personality and Social Psychology 27:41–57.
1980 Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Andrews, F. M.
1984 "Construct validity and error components of survey measures: a structural modeling approach." Public Opinion Quarterly 48:409–442.

Antil, J. H., and P. D. Bennett
1979 "Construction and validation of a scale to measure socially responsible consumption behavior." Pp. 51–68 in K. E. Henion and T. C. Kinnear (eds.), The Conserver Society. Chicago: American Marketing Association.

Bagozzi, R. P.
1981 "Attitudes, intentions, and behavior: a test of some key hypotheses." Journal of Personality and Social Psychology 41:607–627.

Bentler, P. M., and G. Speckhart
1979 "Models of attitude-behavior relations." Psychological Review 86:452–464.
1981 "Attitude's 'cause' behaviors: a structural equation analysis." Journal of Personality and Social Psychology 40:226–238.

Brooker, G.
1976 "The self-actualizing socially conscious consumer." Journal of Consumer Research 3:107–112.

Crosby, L. A., and J. D. Gill
1981 "A causal path analysis of ecological behavior relating to marketing." Pp. 630–636 in K. B. Monroe (ed.), Advances in Consumer Research. Ann Arbor, MI: Association for Consumer Research 8.

Crosby, L. A., J. D. Gill, and J. R. Taylor
1981 "Consumer/voter behavior in the passage of the Michigan container law." Journal of Marketing 45:19–32.

Crosby, L. A., and D. D. Muehling
1982 "External variables and the Fishbein model: mediation, moderation, or direct effects?" Pp. 94–99 in A. A. Mitchell (ed.), Advances in Consumer Research. Ann Arbor, MI: Association for Consumer Research 9.

Crosby, L. A., and J. R. Taylor
1982 "Consumer satisfaction with Michigan's container deposit law—an ecological perspective." Journal of Marketing 46:47–60.
1983 "Psychological commitment and its effects on post-decision evaluation and preference stability among voters." Journal of Consumer Research 9;413–431.

Dunlap, R.
1975 "The impact of political orientations on environment attitudes and actions." Environment and Behavior 7:428–454.

Endo, R., and J. M. Nielsen
1974 "Beyond public opinion: behavioral responses to air pollution." In M. Orleans and G. F. White (eds.), Carbon Monoxide and the People of Denver. Boulder, CO: Institute of Behavioral science.

Farley, J. U., D. R. Lehmann, and M. J. Ryan
1981 "Generalizing from 'imperfect' replication." Journal of Business 54:597–610.

Fishbein, M.
1967 Readings in Attitude Theory and Measurement. New York: John Wiley.
1976 "Extending the extended model: some comments." Pp. 491–497 in B. B. Anderson (ed.), Advances in Consumer Research. Cincinnati, OH: The Association for Consumer Research 3.

Fishbein, M., and I. Ajzen
1975 Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley.
1981 "On construct validity: a critique of Miniard and Cohen's paper." Journal of Experimental Social Psychology 17:340–350.

Henion, K. E.
1976 Ecological Marketing. Columbus, OH: Grid, Inc.

Henion, K. E., and W. H. Wilson
1976 "The ecologically concerned consumer and locus of control." Pp. 131–144 in Ecological Marketing, Proceedings of the First National Workshop on Ecological Marketing. Chicago: American Marketing Association.

Herberger, R. A., and D. T. Buchanan
1971 "The impact of concern of ecological factors on consumer attitudes and buying behavior." Pp. 644–646 in Combined Proceedings. Chicago: American Marketing Association 33.

Jöreskog, K. G., and D. Sörbom
1981 "Lisrel: analysis of linear structural relationships by the method of maximum likelihood, version v." Chicago: National Education Resources, Inc.

Kassarjian, H. H.
1971 "Incorporating ecology into marketing strategy: the case of air pollution." Journal of Marketing 35:61–65.

Kinnear, T. C., and J. R. Taylor
1973 "The effect of ecological concern on brand perception." Journal of Marketing Research 10:191–197.

Kinnear, T. C., J. R. Taylor, and S. A. Ahmed
1974 "Ecologically concerned consumers: who are they?" Journal of Marketing 38: 20–24.

Lepisto, L. R.
1979 "Environmental product attributes and product preference." Pp. 114–117 in R. Franz, R. Hopkins, and A. G. Toma (eds.), Proceedings. Lafayette, LA: Southern Marketing Association.

Loken, B.
1982 "The theory of reasoned action: examination of the sufficiency assumption for a television viewing behavior." Pp. 100–105 in A. A. Mitchell (ed.), Advances in Consumer Research. Ann Arbor, MI: Association for Consumer Research 9.

Maloney, M. P., M. P. Ward, and G. N. Braucht
1975 "A revised scale for the measurement of ecological attitudes and knowledge." American Psychologist 30:787–790.

Mazmanian, D., and P. Sabatier
1981 "Liberalism, environmentalism, and partisanship in public policy-making." Environment and Behavior 13:361–384.

Miniard, P. W., and J. B. Cohen
1979 "Isolating attitudinal and normative influences in behavioral intentions." Journal of Marketing Research 16:102–110.
1981 "An examination of the Fishbein-Ajzen behavioral intention model's concepts and measures." Journal of Experimental Social Psychology 17:309–339.

Murphy, P. E.
1978 "Environmentally concerned consumers: demographic dimensions." Pp. 316–320 in S. Jain (ed.), Proceedings of the 1978 Educator's Conference. Chicago: American Marketing Association.

Page, T. J.
1982 "Comments on investigations of the theory of reasoned action." Pp. 106–108 in A. A. Mitchell (ed.), Advances in Consumer Research. Ann Arbor, MI: Association for Consumer Research 9.

Ryan, M. J.
1982 "Behavioral intention formation: the interdependency of attitudinal and social influence variables." Journal of Consumer Research 9:263–278.

Ryan, M. J., and M. J. Etzel
1976 "The nature of salient outcomes and referents in the extended model." Pp. 485–490 in B. B. Anderson (ed.), Advances in Consumer Research. Cincinnati, OH: Association for Consumer Research 3.

Van Liere, K. D., and R. E. Dunlap
1980 "The social bases of environmental concern: a review of hypotheses, explanations, and empirical evidence." Public Opinion Quarterly 44:181–197.
1982 "Environmental concern: does it make a difference how it's measured?" Environment and Behavior 13:651–676.

Webster, F. E.
1975 "Determining the characteristics of the socially conscious consumer." Journal of Consumer Research 2:188–196.

Wheaton, B., B. Muthén, D. F. Alwin, and G. F. Summers
1977 "Assessing reliability and stability in panel models." Sociological Methodology 84–136.

Change and Stability in Presidential Popularity at the Individual Level

KENT L. TEDIN

THE STANDARD approach to the study of change in presidential popularity has been to plot survey observations over time and treat each observation as an independent data point. This approach has led to a focus on macro-level predictors and the use of econometric (time series) statistical models. Change at the micro level, on the other hand, has received far less attention (for an exception, see Wanat, 1982).

However, from the standpoint of both the practice and theory of democracy the quality of individual change is of considerable importance. It is commonly agreed that presidential popularity is an important ingredient in the mix of attributes that constitute presidential power—the ability of the president to get things done (Neustadt, 1973). Are those persons who affect this power by altering their opinions located on the outskirts of political life, or does change occur among the most interested and sophisticated segments of society, or is it evenly spread? Evidence from the analysis of issue positions taken by panel respondents leads us to expect the latter. Most investigations into the reliability of the mass response on repeated measures show a stable attitude and measurement error, with the error spread evenly across levels of sophistication (Achen, 1975; Judd and Milburn, 1980).

Abstract This article employs a three-wave panel to examine change and stability in the popularity of the president. By survey standards, the item is found to be quite reliable, approaching the reliability of party identification. Change in the item can be predicted using measures of political interest and sophistication, partisanship, and strength of presidential approval or disapproval.

Kent L. Tedin is Professor and Chairman of the Department of Political Science, University of Houston.

Public Opinion Quarterly Vol. 50:555–562 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-555/$2.50

Also important for the interpretation of presidential popularity is the behavior of the "undecided" respondents. Mueller (1973) argues that presidential approval is the mirror image of disapproval. Surges and declines do not come from the undecideds, but are the product of radical change. The undecided category is not a stopover for those who are in the process of changing their mind. Wanat's (1982) research supports Mueller on this point—at least with regard to international rally points. On the basis of a method for estimating cell frequencies from table marginals he observed that radical change was substantially more frequent than change involving the undecided category. He concluded that undecideds "are largely irrelevant when change in aggregate approval is examined. They are not participants in political life and may be ignored for some purposes" (1982:187). This statement implies a more or less stable base of undecided respondents, with little movement in and out of the category. A micro-level analysis of the behavior of undecideds should tell us something about the implications of basing presidential popularity on all respondents (as is usually the case in descriptive studies) or on only those with a negative or positive evaluation (more commonly seen in analytic pieces where popularity is operationalized by subtracting the negatives from the positives and ignoring the undecideds; see Stimson (1976) for an example of this approach).

Data

Data for this study were taken from a three-wave panel conducted in Texas. The first wave was executed between March 7 and March 15, 1982. A sample was drawn from the list of registered voters and stratified along racial, income, regional and urban-suburban lines to make it representative of the state's electorate. A total of 750 respondents was interviewed by telephone on the first wave. The second wave was conducted between May 10 and May 17 and 610 of the original 750 were reinterviewed. The final wave was executed between November 7 and November 15, yielding a total of 531 usable responses.

President Reagan's average 1982 popularity over the three waves of our panel was 56 percent, with 58 percent approving in March, 55 percent in May, and 54 percent in November. Given the absence of any event which might have had a substantial impact on presidential popularity, our investigation into the properties of the item must be placed within the context of little aggregate change. We shall address the following questions: (1) How reliable was the measurement of presidential popularity? (2) To what extent was change of the "radical" variety as opposed to movement in and out of the undecided category? (3) What was the pattern of micro-level change? (4) What variables predict who changes and who remains stable?

Analysis

There are now numerous studies which attempt to account for the largely unpatterned individual-level change which characterizes most opinions toward public issues. The developing consensus (the "latent attitude" model) holds that while the indicators may show considerable fluctuation over time, much of the public has reasonably well-formed underlying attitudes. But questions of latent attitudes aside, in the case of presidential popularity it is the behavior of the *indicator* that gets reported by the press and read by the public—almost all of whom have no technical knowledge concerning opinion measurement. Panel analyses have shown single-item indicators to be notoriously bad measures of underlying concepts (Inglehart, 1985). Thus, in the peculiar instance of support for the president, the quality of measurement has unique importance.

THE RELIABILITY OF PRESIDENTIAL POPULARITY

Presented in Table 1 are the test-retest correlations for presidential popularity both with and without the undecideds. Also presented are the test-retest correlations for party identification, liberal–conservative ideology, and the governor's rating. Somewhat over one-half the variation in presidential popularity is explained by the previous measurement. This figure is not as high as for party identification, but is considerably higher than for ideology or the evaluation of the governor (whose aggregate approval was virtually unchanged over the panel—49 percent, 43 percent, and 46 percent).

With almost one-half the test-retest variation unexplained, it is apparent that presidential popularity is characterized by considerable movement at the individual level. Nevertheless, by survey standards, the item is reliably measured—a point of some consequence, given the practical importance of the statistic. It has a constant reliability of .74, which compares favorably to party identification which has a reliability of .80. On the other hand, the reliability of political ideology was only .51 and the reliability of the governor's rating was only slightly higher at .52.

ACCOUNTING FOR CHANGE

The recent literature on the stability of attitude indicators would suggest that much of the unexplained variation in the test-retest corre-

Table 1. Test-Retest Product-Moment Correlations for Texas Panel

	Texas Panel			
	March-May	*May-Nov.*	*March-Nov.*	*Reliability*[a]
Pres. popularity[b]	.74(531)	.67(531)	.69(531)	.74
Pres. popularity[c]	.81(469)	.74(445)	.73(468)	.82
Party identification	.83(525)	.82(523)	.85(524)	.80
Liberal/conservative[d]	.62(486)	.48(487)	.57(486)	.51
Rate Gov. Clements[e]	.53(531)	.51(531)	.52(531)	.52

[a] The constant reliability estimate is $(r_{12})(r_{23})/(r_{13})$. A separate reliability estimate (Wiley and Wiley, 1970) yields the same inference.

[b] Categories are approve, undecided, disapprove.

[c] Categories are approve and disapprove only.

[d] Respondents were asked to place themselves on a 1–7 scale.

[e] Respondents were asked to rate the governor on a 0–10 scale.

lations for presidential popularity is random. We shall investigate the properties of change by first examining the 27 possible combinations of approve-undecided-disapprove over the three waves of the panel (presented in Table 2). Keeping in mind the context of very little aggregate change, we find that three patterns are apparent. First, the most frequent responses were approve-approve-approve and the converse. Stable respondents accounted for 67 percent of the sample. Second, only two of the 531 respondents were undecided in all three waves. Thus, the inert group is very small. Third, there was much more movement in and out of the undecided category than past research would lead us to expect. In fact, radical change (movement from approve to disapprove and the reverse) was only about three-quarters as frequent as movement in and out of the undecided category. A majority of those who changed their response did not go the entire distance to a complete reversal. While no more than 12 percent indicated that they were undecided at one point in the panel, 22 percent were undecided at some point during the seven-month period.

An "eyeballing" of Table 2 does not reveal any obvious pattern to the change process. No one particular change permutation attracts a greatly disproportionate number of respondents. In fact, the arrangement looks very much like what we have come to expect in panel studies of public opinion—a stable attitude and measurement error. A series of standard computations (Converse, 1964; Wiley and Wiley, 1970) testing for patterned vs. random change yielded mixed results.[1]

---

[1] One test is to assume that true change can be described by a Markovian process, and that observations at time 1 influence those at time 3 only through time 2. This yields the prediction that $r_{13} = (r_{12})(r_{23})$. The former coefficient is .69; the latter is .50. Using the LISREL measurement model, which for single items reduces to the Wiley and Wiley (1970) method, yields an inference of a stable attitude and measurement error. However, I am

Table 2. Change in Presidential Popularity over Three Time Points ($N = 531$)

Responses[a]	*Number of Cases*	*Percent*
A-A-A	237	44.6%
A-A-D	11	2.1
A-D-A	10	1.9
A-D-D	8	1.5
A-A-U	10	1.9
A-U-D	4	0.7
A-U-U	2	0.3
A-U-A	21	3.9
A-D-U	0	0.0
D-D-D	117	22.0
D-A-A	9	1.7
D-D-A	18	3.4
D-A-D	9	1.7
D-U-A	8	1.5
D-D-U	5	0.9
D-U-D	16	3.0
D-U-U	0	0.0
D-A-U	0	0.0
U-A-A	7	1.3
U-A-D	2	0.3
U-D-A	3	0.6
U-U-U	2	0.4
U-U-A	6	1.1
U-D-U	8	1.5
U-D-D	12	2.2
U-U-D	2	0.4
U-A-U	1	0.2

[a] A = approve; U = undecided; D = disapprove.

However, a close look at Table 2 shows the change permutations are not equaprobable as would be required by a model positing a stable attitude and random error (Inglehart, 1985). For example, among those unstable at T2-T3 the modal response at T3 is to *move back* to the evaluation held at T1 (e.g., approve-undecided-approve). Forty-seven percent of the unstable at T2-T3 show this nonrandom process.[2]

We shall next address the question of whether certain types of respondents are more likely to remain stable in support of the president

---

dubious that assumption of no serial correlation in the error terms is being met. On the other hand, if all change is random then for the T1-T2 shifters (for example) their opinion at T2 should not predict their opinion at T3. However, some of the permutations for this exercise yield statistically significant product-moment correlations. Also, if all change is random then the presidential evaluation of the shifters should not be related to issue questions beyond that expected by chance. Again, we find numerous statistically significant relationships (although the relationships for stable subgroup are considerably higher than for the unstable subgroup).

[2] An alternative explanation might be common response set in the independent and dependent variables or some other type of systematic response error.

than are others. The focus here is on stability, as any attempt to tackle the methodological pitfalls involved in separating true change from measurement error is beyond the scope of this article.

The dependent variable for this analysis is a binary scoring of respondents, over the three waves of the panel, as stable or unstable in their evaluations of the president. Three types of independent variables will be used: (1) political interest and sophistication, (2) partisanship, and (3) strength of presidential approval or disapproval. The null hypothesis of no relationship is taken from the recent literature on attitude stability (Achen, 1975; Judd and Milburn, 1980) which states that change in panel data is largely due to faulty measurement and is therefore spread evenly across subgroups. The alternative hypothesis (Converse, 1964) holds that for those most psychologically involved in politics the evaluation of the president is part of an integrated belief system. These individuals cling more firmly to their opinion because it has a better defined evaluative and/or cognitive support structure. Consequently, significant relationships should exist between the independent variables and stability. Of course, the test of this hypothesis also has implications for the interpretation of unreliability—i.e., to what extent is it due to the instrument vs. the individual (see Achen, 1975 for an elaboration).

The data clearly support the alternative hypothesis. The bivariate relationship (gamma) between the dependent variable and the political sophistication independent variables is .26 for knowledge of public figures, .23 for newspaper readership, .32 for the number problems mentioned by respondents to an open-ended query about the problems facing the state, and .28 for education (all, as well as those reported below, show a X^2 statistic significant beyond .05 unless otherwise mentioned). Political interest and activity display a similar pattern. Respondents were asked if they voted in the spring, 1982 primaries (38 percent said yes), and were also asked, on a one-to-four scale, to indicate the level of their political interest. The relationship is .33 for the former and .24 for the latter. The strength with which the attitude is held also predicts stability. For respondents who answered "approve" or "disapprove" a follow-up probe asked if the person "strongly" or "somewhat" approved/disapproved.[3] Given three waves of the panel, this variable can range from zero to three (the undecideds were excluded). Gamma is .41, indicating that those who felt strongest were most stable over the seven-month period.

---

[3] This correlation is based on 422 respondents who did not, at some point, answer "undecided." Including the undecideds would make the dependent variable partially tautological to the independent variable.

Table 3. Cross-tabulation of Party Identification and Stability of Presidential Popularity (N = 519)

	Partisanship						
	Stg Dem	*Weak Dem*	*Ind Dem*	*Ind*	*Ind Rep*	*Weak Rep*	*Stg Rep*
Stable approve	04%	23%	36%	55%	79%	85%	90%
Stable disapprove	55	30	19	09	05	15	10
Subtotal stable	59%	53%	55%	64%	84%	85%	90%
Unstable	41	47	45	36	16	15	10
	100%	100%	100%	100%	100%	100%	100%
	(138)	(88)	(68)	(49)	(77)	(40)	(59)

In terms of partisanship, it was the Democrats who were most unstable in their presidential evaluation. On the other hand, 90 percent of the strong Republicans and 85 percent of the weak Republicans were stable. That stability was almost entirely approve-approve-approve (see Table 3). Even the Independents were more stable than the Democrats—again with stable approvers accounting for the bulk of that stability. Only among strong Democrats did stable disapprovers greatly outnumber stable approvers. The lack of stability among Democrats may be due to the cross-pressures experienced by their white, conservative adherents. They are pulled in one direction by their partisanship and in another by their political ideology. The manifestation of this cross-pressure was opinion instability—44 percent of the white, conservative Democrats were unstable. On the other hand, 70 percent of the blacks were stable (all D-D-D). Of course, the political cues they receive tend to point overwhelmingly in one direction.

Conclusion

Though it is but a single question, the standard presidential popularity item shows a test-retest reliability that compares favorably with the best measured political attitudes. Also, the modest but consequential relationships between the independent variables and stability suggests that instrument problems are not as significant a cause of unreliability as is the case with issues. Much of the movement in the panel was due to respondents who at some point were undecided. They outnumbered those displaying "radical" change by about 1.5 to 1. The analysis showed generally that change during the period of this study was most prevalent among those least involved in politics. It seems reasonable to infer, therefore, that one group likely to rally to the president in time of crisis or withdraw support when the news is bad tends to be those on the fringes of political life. The irony that Converse (1962) noted with regard to floating voters and media attention seems also to hold with regard to change in presidential popularity. That is, strong partisans and the most politically attentive are, in general, the least likely to change.

References

Achen, Christopher H.
1975 "Mass political attitudes and the survey response." American Political Science Review 69:1218–1231.
Converse, Philip
1962 "Information flow and the stability of partisan attitudes." Public Opinion Quarterly 26: 578–99.
1964 "The nature of belief systems in mass publics." In David E. Apter (ed.), Ideology and Discontent. New York: The Free Press.
Inglehart, Ronald
1985 "Aggregate stability and individual-level flux in mass belief systems: the level of analysis paradox." American Political Science Review 79: 97–116.
Judd, Charles, and Michael Milburn
1980 "The structure of attitude systems in the general public." American Sociological Review 45: 627–43.
Mueller, John E.
1973 War, Presidents, and Public Opinion. New York: Wiley and Sons.
Neustadt, Richard
1973 Presidential Power. New York: Wiley and Sons.
Stimson, James A.
1976 "Public support for American presidents: a cyclical model." Public Opinion Quarterly 40: 1–21.
Wanat, John
1982 "The dynamics of presidential popularity shifts: estimating the degree of opinion shift from aggregate data." American Politics Quarterly 10: 181–89.
Wiley, David, and John Wiley
1970 "The estimation of measurement error in panel data." In Hubert M. Blalock (ed.), Causal Models in the Social Sciences. Chicago: Aldine.

Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos

STEPHEN D. REESE, WAYNE A. DANIELSON,
PAMELA J. SHOEMAKER, TSAN-KUO CHANG,
AND HUEI-LING HSU

SURVEY researchers have long been aware that interviewers can affect the very responses they are assigned to gather. Fortunately, the variation in items attributable to who asks the questions is generally quite small (Tucker, 1983). However, race is one interviewer characteristic which exerts a systematic and significant biasing effect, particularly on answers to race-related questions. This study examines this "ethnicity"-of-interviewer effect in a survey of both Mexican-Americans and Anglos. As the fastest growing minority in the U.S., Hispanics have attracted increasing research interest. To reach them, bilingual interviewers have been used (e.g., Greenberg et al., 1983), making potential interviewer effects a practical methodological issue. In addition to posing a validity concern, Schuman and Converse (1971) note that interviewer effects represent a useful indicator of relations between the racial (and ethnic) groups, a "fact of social life."

The current thinking is that interviewer effects can be explained through a theory of general deference. Respondents try to avoid offending an interviewer, specifically on items which mention the inter-

---

Abstract A general interpersonal deference explanation for ethnicity-of-interviewer effects is tested with a random survey of Anglos and Mexican-Americans using Anglo and Hispanic interviewers and questions related to Mexican-American culture. The theory was most clearly supported among Anglo respondents. The mixed support found among Mexican-Americans suggests that deference is strongest when questions deal with the culture represented by the interviewer, not the respondent.

Stephen Reese and Pamela Shoemaker are Assistant Professors in the Department of Journalism at the University of Texas at Austin, 78712, where Wayne Danielson is Jesse H. Jones Professor and Tsan-Kuo Chang is a doctoral candidate. Huei-Ling Hsu is a doctoral student at the Institute for Communication Research at Stanford University.

Public Opinion Quarterly Vol. 50:563–572 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press
0033-362X/86/0050-563/$2.50

viewer's race. Most studies supporting such a race-of-interviewer effect, however, have examined only blacks and whites. Hyman et al. (1954) found that blacks gave different responses to whites than to other blacks, presumably out of deference to the dominant "caste." Black responses to items regarding militant protest and hostility to whites are particularly sensitive to race of interviewer, while blacks' reports of discrimination, living conditions, and personal background are not (Schuman and Converse, 1971). Subsequent studies have supported a more general interpersonal deference theory, which accounts for interviewer effects on whites as well as blacks (Athey et al., 1960, Hatchett and Schuman, 1975–76, and Campbell, 1981). What's more, a face-to-face interview is not required; the race-of-interviewer effect in a telephone interview has been found which is similar in magnitude to that reported from a personal survey (Cotter et al., 1982). This finding is of particular interest, given the predominance of telephone surveys as a data-gathering method.

Other ethnic groups have also been examined for interviewer effects, but the focus of these studies has been on nonracially sensitive items. Weeks and Moore (1981) found that a difference in ethnicity (Cuban, Chicano, Puerto Rican, and Chinese) between interviewers and respondents did not affect responses. Similarly, Welch et al. (1973) reported that Mexican-Americans did not differ in their responses to Anglos and other Mexican-Americans. Both studies were based on personal interviews about general items, supporting the basic theory that general deference does not occur for questions which aren't related to ethnicity.

This study tests the general deference theory using a telephone survey of Mexican-Americans and Anglos, with both Hispanic and Anglo interviewers. Both general and ethnicity-related questionnaire items are examined, including questions not previously studied for interviewer effects—ethnic media behavior and ethnic environment. Possible bias in such items is particularly important, given their frequent use in measures of acculturation.

On the basis of previous research it is hypothesized that no interviewer effects will be found for nonethnicity-related items (e.g., education), but that both Anglos and Mexican-Americans will defer to opposite-ethnicity interviewers on questions related to Mexican-American culture. In addition, it is hypothesized that among Mexican-Americans this deference will extend to their self-reports of ethnic environment and behavior. In other words, according to the general deference theory we would expect respondents to overreport ethnic behavior when interviewed by someone from the same ethnic background, and underreport it when interviewed by someone else.

Methods

SAMPLE

Between February and May of 1984, telephone interviews were conducted with 1076 Texas residents over the age of 18, using a random-digit dialing technique.[1] In order to supplement the number of Mexican-American respondents, a special sample of persons with Hispanic surnames was selected. This random sample was drawn using telephone books from cities with large Hispanic populations.[2] An additional 142 interviews were added to the data using this method. The random-digit sample produced a completion rate of 55.7 percent and a refusal rate of 21 percent. The supplementary sample yielded a completion rate of 69.6 percent and a refusal rate of 8.8 percent.[3] In both samples, interviewers asked for a male or female resident at random and made as many as five callbacks on different days to reach a respondent. Only after five unsuccessful attempts or if the number was nonworking or a business was it replaced with another from the cluster.

Both Hispanic and Anglo bilingual interviewers were hired to administer the questionnaires. Each interviewer spoke fluent regional Spanish and unaccented English. If a respondent said he or she spoke some Spanish, the interviewer offered to do the questionnaire in Spanish or to clarify any words or phrases in Spanish. The first question put to respondents was which language they spoke with family and friends.

---

[1] Two hundred telephone exchanges were drawn at random without replacement from a list of all state exchanges. A working "seed" bank within each exchange was established during pretesting, and 20 random two-digit numbers were attached to each seed number (e.g., (512) 345–13——) to create the numbers in each of 200 "clusters." Four interviews were attempted in each cluster for a target sample size of 800. Those exchanges yielding two or more Hispanic interviews were oversampled to reach the final *N* of 1076.

[2] These areas included Laredo, Encinal, San Antonio, Corpus Christi, Alamo, Brownsville, Donna, Edcouch, Edinburg, Elsa, Hargill, Harlingen, Hidalgo, La Feria, Los Fresnos, Lyford, McAllen, Mercedes, Mission, Pharr, Port Isabel, Raymondville, Rio Hondo, San Benito, San Juan, Santa Rosa, Sullivan City, Weslaco, and El Paso. Numbers were randomly chosen from each community's listings, proportionate to its size, such that towns with larger populations contributed more respondents than others. The selection of telephone numbers was conducted by first choosing a random page, column, and location and then selecting the nearest number associated with an Hispanic surname.

[3] The completion rates were calculated as follows: (completed interviews)/[(total numbers called) – (numbers out of service + numbers changed + business numbers + noneligible respondents)]. The total numbers called included "no answer" after five tries. The refusal rates were the proportion of those contacted who refused the interview. The distribution of telephone numbers yielding completed interviews reflects the actual distribution of numbers in exchanges across the state, and the demographic characteristics of respondents are consistent with 1980 state census parameters. A separate study by the authors showed that the two methods yielded samples which can be considered to represent the same population of Hispanics (Chang et al., 1985).

If some Spanish was spoken they were given the opportunity to be interviewed in Spanish. A total of 86 questionnaires were conducted in Spanish. Ethnicity of interviewer affected neither respondents' self-reported language nor the language spoken during the interview. Only those respondents questioned by interviewers who had either clearly Anglo or Hispanic surnames and who used their names in the questionnaire introduction were included in the analysis. Questionnaires were randomly distributed among these individuals, such that each interviewer conducted a cross-section of interviews from around the state. This procedure yielded 1004 total interviews for analysis. Of this total, Hispanics conducted 724 interviews, 199 (27.5 percent) of them with other Hispanics. Anglos administered interviews, 85 (30.4 percent) with Hispanics.[4] As reflected in these proportions, no significant differences in response rates were found for the two groups of interviewers.

MEASURES

Respondents were considered Mexican if they said their ethnic background was wholly or partially Hispanic, which included Mexican-Americans, Puerto Ricans, and Cubans. Mexican-Americans accounted for most of the Hispanic respondents (86.4 percent).[5] The culturally related items included ethnic media use, attitudes toward Mexican-Americans, attitudes toward bilingualism, and reports of ethnic interpersonal environment. Exact questionnaire items are included in the Appendix.

Results

The results support the deference theory most clearly among the Anglo respondents. As seen in Table 1, four of the six comparisons show statistically significant differences. Anglos were less likely to

---

[4] The data examined in this study were part of a larger project which examined a cross-section of Texas residents, with a particular focus on the media behavior, political attitudes, and ethnicity of Hispanics. The total questionnaire took about 20 minutes to administer. The research was funded by a Gannett foundation grant to the first three authors.

[5] The label *Mexican-American* is used throughout this study because it is the most common name applied to Hispanics in Texas. Many of the cultural items in the questionnaire refer specifically to "Mexican-Americans." Because of the common Spanish heritage, it is thought that the few non-Mexican-American Hispanics in this study would respond similarly in the case of the race-of-interviewer effect. By the same token, the few blacks (10 percent of the total) and other non-Hispanic ethnic respondents were included in the Anglo category for purposes of discussion because the theoretical distinction is between those interviewed by someone from the same ethnic background and by someone of a different ethnicity. Blacks, for example, would be expected to behave the same as whites when interviewed by an Hispanic.

Table 1. Ethnicity-of-Interviewer Effects on Mexican-American and Anglo Respondents: Cultural Evaluations and Issues

		Anglo Respondents		*Mexican-American Respondents*	
		Anglo Interviewer	*Hispanic Interviewer*	*Anglo Interviewer*	*Hispanic Interviewer*
N Interviewers/*N* Respondents[a]		5/164	12/453	4/85	11/199
Prefer Mexican-American culture and customs to others	Agree	9.4%	23.2%[b]	58.1%	58.2%
	No opinion	19.3	30.0	26.0	18.5
	Disagree	71.3	46.8	15.9	23.3
Would rather associate with Mexican-Americans than with others	Agree	8.2	9.7	42.3	34.4
	No opinion	35.6	37.9	25.7	17.9
	Disagree	56.2	52.4	32.0	47.7
Mexican-Americans are a lot different from other people	Agree	31.2	19.8[b]	41.5	25.8
	No opinion	19.8	14.7	2.8	12.4
	Disagree	49.0	65.5	55.7	61.8
Everyone should be able to speak Spanish	Agree	26.7	31.9[b]	57.9	60.4
	No opinion	10.4	14.5	8.2	7.3
	Disagree	62.9	53.6	33.9	32.3
Grade school classes should be taught in Spanish and English	Agree	37.8	41.5[b]	73.7	67.1
	No opinion	6.1	9.1	3.7	7.4
	Disagree	56.1	49.4	22.6	25.5
Everyone should be able to speak English	Agree	82.5	81.6	80.7	84.5
	No opinion	8.1	7.0	3.6	7.3
	Disagree	9.4	11.4	15.7	8.2

[a] Although interviewer ethnicity is discussed as a characteristic of respondents, statistical tests are based on the *N* of interviewers, yielding an approximation of results obtained using the respondent as the unit of analysis, but a more conservative statistical test. Values represent the mean proportions of respondents selecting each response, averaged across interviewers.

[b] Differences are statistically significant based on a one-tailed *t*-test (df = 15), $p < .05$.

disagree that they preferred Mexican-American culture when interviewed by a Hispanic rather than by another Anglo. Note that the *N* of interviewers is used for statistical tests, an approach used by, for example, Singer et al. (1983).

Similarly, Anglos were more likely to disagree that Mexican-Americans are a lot different from other people when asked by a Hispanic. Anglos were less favorable toward bilingualism when interviewed by an Anglo; they were more likely to disagree that everyone should be able to speak Spanish and that grade school classes should be taught in both English and Spanish when interviewed by an Anglo than by a Hispanic. Statistically significant results were not found, however, for the same items among Mexican-American respondents, and, in fact, the responses shown in Table 1 show effects in two directions. For two of the three cultural evaluation items, there appears to be a reverse deference effect. When interviewed by an Anglo rather than by another Hispanic, Mexican-Americans were more likely to agree that they would rather associate with other Mexican-Americans and that they are a lot different from others. For the language attitude items, although Hispanics were more likely to agree that grade school classes should be bilingual when asked by an Anglo than when asked by another Hispanic, responses to the other two items showed very slight differences in the other direction.

Other items were examined which dealt specifically with ethnicity, so only Mexican-Americans' responses were included in Table 2. Again, results are found both for and against the deference hypothesis. Three items are supportive (one statistically significant): Hispanics reported watching more TV news in Spanish and having more Mexican-American friends and neighbors when interviewed by another Hispanic than by an Anglo. On the other hand, they reported reading less in Spanish and perceived having a Spanish-language newspaper as less important when interviewed by a Hispanic than when interviewed by an Anglo.

Other noncultural questionnaire items were also examined for interviewer effects (not shown here). The only ones which produced substantial differences between Hispanic and Anglo interviewers (although not statistically significant) were education and income. When interviewed by an Anglo, Mexican-Americans reported higher educational attainment and higher family income than when interviewed by a Hispanic.

Although no interviewer effects were expected for ethnicity items (in Table 2) among Anglos (given that so few reported ethnic behavior), two statistically significant differences were found supporting the deference hypothesis. Every Anglo (100 percent) interviewed by another Anglo said having a Spanish paper to read was not important, while 96 percent told Hispanic interviewers it was not important. Similarly, 99 percent of Anglos told Anglo interviewers that half or fewer of their friends were Mexican-American while 96 percent gave those responses to Hispanic interviewers.[6] Other analyses were conducted for other questionnaire items, but no significant differences were found and the results are not reported. For example, respondents were asked to evaluate various Hispanic political groups, and to indicate how well the media were doing in reporting on "people like them." These ambiguous reference items and the appraisals of specific groups were apparently not subject to the same interviewer effects found for more personally direct items relating to cultural behavior and attitudes. Respondents were also asked about their general media use and at-

---

[6] For the items in this study where significant differences were found, calculating Eta^2—the proportion of variance in responses attributable to interviewer ethnicity—produced values between 1 and 3 percent, comparable to the size of effects found in previous studies.

Table 2. Ethnicity-of-Interviewer Effects on Mexican-Americans: Ethnic Media, Interpersonal Environment, and Socioeconomic Status

	Anglo Interviewer	*Hispanic Interviewer*
N Interviewers/*N* Respondents[a]	4/83	11/199
How much TV news watched in Spanish		
None	73.8%	65.7%
At least some	26.2	34.3
Overall, how much TV watched in Spanish		
None	46.2	45.5
At least some	53.8	54.5
How much radio listened to in Spanish		
None	41.0	42.2
At least some	59.0	57.8
How many newspapers and magazines read in Spanish		
None	67.5	85.3[b]
At least some	32.5	14.7
Importance of having daily newspaper to read in Spanish		
Not important	42.7	49.6
At least a little important	57.3	50.4
How many Mexican-American friends		
Half or less	62.2	38.7
More than half	37.8	61.3
How many Mexican-American Neighbors		
Half or less	69.1	42.2[b]
More than half	30.9	57.8

[a] Although interviewer ethnicity is discussed as a characteristic of respondents, statistical tests are based on the *N* of interviewers, yielding an approximation of results obtained using the respondent as the unit of analysis but a more conservative statistical test. Values represent the mean proportions of respondents selecting each response, averaged across interviewers.

[b] Differences are statistically significant based on a one-tailed t-test ($df = 14$), $p < .05$.

titudes toward the (English language) media. A complete list of the questionnaire items is available from any of the first three authors.

Discussion

This study shows that an ethnicity-of-interviewer effect exists for both Mexican-Americans and Anglos in a telephone survey context. The results partially confirm hypotheses derived from a theory of interpersonal deference. Even presumably more "objective" self-reports of behavior and conditions are subject to interviewer effects as are the more "subjective" attitudinal items (although not always in the predicted direction).

The responses by Anglos to the cultural evaluation and bilingualism items are most compatible with the deference theory. Those questions may be considered the most potentially sensitive to an Hispanic interviewer and, indeed, showed the strongest effects. However, those same items failed to show interviewer effects among Hispanics themselves. The latter results are nevertheless consistent with those reported by Cotter et al. (1982), who found a general lack of interviewer effects among blacks when the questions were about blacks. The authors suggested that blacks felt no need to alter responses about their own race but that deference effects would be found if blacks were asked about whites. We would expect similar findings in this study had Mexican-Americans been asked about Anglos.

Other items showed mixed results and may require different explanations beyond deference. In the cases where Hispanics exaggerated their ethnic behavior (TV news watched in Spanish) and proportion of ethnic contacts (friends and neighbors), a cultural affinity or "halo" effect may be at work. This phenomenon represents, perhaps, an effort to appear more like the ethnic group represented by the interviewer, which calls for either enhancing or diminishing reported ethnic characteristics, depending on who asks the questions. Although not overtly cultural, education and income questions may also be said to produce responses resulting from this process. Mexican-Americans are no doubt aware of the socioeconomic disparity between themselves and Anglo society and may shade their status to more closely match the perceived status of the interviewer.

Conversely, two items related to Spanish-language print media use produced results contrary to this explanation. Reading and perceived importance of Spanish-language print media were greater when Mexican-Americans responded to Anglo interviewers than when they responded to other Hispanics. Further study is needed to examine which types of ethnicity-related items show a cultural affinity effect and which a reverse affinity effect, and why.

The results discussed above, although based on an *N* of 1004, used the interviewer as the unit of analysis, thus producing a more conservative test than when the number of respondents is the unit of analysis. Additional research employing a greater number of interviewers than available here is needed before stronger conclusions are drawn, particularly regarding the Mexican-American respondents. What is more, to more fully test the deference theory among Mexican-Americans, other questions need to be examined. Specifically, items should be included which ask Hispanics to evaluate Anglos and Anglo culture, so as to better approximate classic studies of blacks and whites which originally generated the interpersonal deference explanations.

As the Spanish population in the U.S. continues to grow, it will become even more important for surveys to include them in order to be representative. To do so, bilingual interviewers will no doubt be used. This study shows that ethnicity of interviewers can exert a significant biasing effect on a variety of cultural questions, even over the phone.[7] Fortunately, this effect is not large—respondents were not prompted to fundamentally alter their attitudes (from agree to disagree, for example), but they did shade their responses depending on who they talked with. Whether this bias results from general deference, as other work has suggested, or from other processes, such as cultural affinity, is left to future research.

Appendix: Questionnaire Items

1. About how much of the *news* you watch on television is in Spanish? Would you say—none of it, less than half, about half, more than half, all of it?
2. *Overall*, about how much of the television you watch is in Spanish? Would you say—none of it, less than half, about half, more than half, all of it?
3. How much of the time that you listen to the radio is it in Spanish? None of the time, less than half, about half, more than half, all or most of the time?
4. About how many of the newspapers and magazines you read are in Spanish? Would you say—none, less than half, about half, more than half, all of them?
5. How important is it for you to have a daily newspaper to read in Spanish? Would you say—not important at all, a little important, somewhat important, very important?
6. About how many of your neighbors would you say are Mexican-Americans? Almost all, about ⅔, about ½, about ⅓, very few, none?
7. How many of your friends are Mexican-American? Almost all, about ⅔, about ½, about ⅓, very few, none?
8. Would you agree or disagree with the following statements? I prefer Mexican-American culture and customs to other culture and customs. Agree strongly, agree, no opinion, disagree, disagree strongly (the following items shared the same scale).
9. Given the choice, I would rather associate with Mexican-Americans than with other people.
10. Mexican-Americans are a lot different from other people.
11. Everybody who lives in Texas ought to be able to speak English.
12. Everybody who lives in Texas ought to be able to speak Spanish.
13. Grade school classes in Texas ought to be taught in both English and Spanish.

---

[7] In this study the interviewers' names, and slight accent in some cases, were presumed to inform respondents of interviewer ethnicity. Unfortunately, there were too few respondents questioned by "nameless" interviewers for reliable comparisons. It would be informative in future studies to determine if concealing interviewer ethnicity reduces its effect on response bias.

References

Athey, K. R., J. E. Coleman, A. P. Reitman, and J. Tang
1960 "Two experiments showing the effect of the interviewer's racial background on responses to questionnaires concerning racial issues." Journal of Applied Psychology 44:244–46.

Campbell, B. A.
1981 "Race-of-interviewer effects among southern adolescents." Public Opinion Quarterly 45:231–44.

Chang, T. K., W. Danielson, S. Reese, and P. Shoemaker
1985 "Mexican-Americans as an ethnic sample: random digit dialing vs. sampling by surname." Paper presented to the annual conferences of the Association for Education in Journalism and Mass Communication, Memphis, Tennessee.

Cotter, P. R., J. Cohen, and P. B. Coulter
1982 "Race-of-interviewer effects in telephone interviews." Public Opinion Quarterly 46:278–86.

Greenberg, B. S., M. Burgoon, J. K. Burgoon, and F. Korzenny
1983 Mexican Americans and the mass media. Norwood, N.J.: Ablex.

Hatchett, S., and H. Schuman
1975–1976 "White respondents and race-of-interviewer effects." Public Opinion Quarterly 39:523–28.

Hyman, H., et al.
1954 Interviewing in Social Research. Chicago: University of Chicago Press.

Schuman, H., and J. M. Converse
1971 "The effects of black and white interviewers on black responses in 1968." Public Opinion Quarterly 35:44–68.

Singer, E., M. R. Frankel, and M. B. Glassman
1983 "The effect of interviewer characteristics and expectations on response." Public Opinion Quarterly 47:68–83.

Tucker, C.
1983 "Interviewer effects in telephone surveys." Public Opinion Quarterly 47:84–95.

Weeks, M. F., and R. P. Moore
1981 "Ethnicity-of-interviewer effects on ethnic respondents." Public Opinion Quarterly 45:245–49.

Welch, S., J. Comer, and M. Steinman
1973 "Interviewing in a Mexican-American community: an investigation of some potential sources of response bias." Public Opinion Quarterly 37:115–26.

The Polls:
The Most Admired Man and Woman

TOM W. SMITH

> Heroes are created by popular demand, sometimes out of the scantiest materials . . .
>
> —Gerald White Johnson, *American Heroes and Hero-Worship*
>
> Lives of great men oft remind us/we should make our own sublime.
>
> —Henry Wadsworth Longfellow, "A Psalm of Life"

Society's heroes both reflect social norms and shape them. When people select heroes, they choose individuals who exemplify and personify virtues valued by the society. In turn, the heroes become standards of excellence that people can identify with and, if they wish, strive to emulate. By studying those whom society admires we can not only learn about its cast of heroes, but also draw inferences about the values shared by its people. Similarly, by following changes in its hero types we can track how society is changing.

Almost yearly since 1947 Gallup has been asking the American public what man and what woman "living today in any part of the *world*" they admire the most. (For details on wordings and occurrences see the Appendix). These questions comprise the longest, most consistent, and richest data series on the heroes of recent American society.

As with any single indicator, the most-admired question has its strengths and weaknesses. First, since the question is open-ended, we get an essentially unrestricted outpouring of nominations. The generated list is not constrained by a questioner's choices and the instrument is sensitive to changes in public taste, dramatic events, and other changes. The open-ended format has drawbacks, however. In most surveys many minor (often solitary) selections were not preserved,

Tom W. Smith is Senior Study Director, National Opinion Research Centner (NORC), University of Chicago.

Public Opinion Quarterly Vol. 50:573–583 © 1986 by the American Association for Public Opinion Research
Published by The University of Chicago Press 0033-362X/86/0050-573/$2.50

being lumped together into a nameless miscellany. In addition, open-ended questions require respondents to think more than closed alternative questions, and as a result, item nonresponse is often quite high. Second, until 1961 only the first choice of each respondent is recoded so it is not possible to study within-subject variation over the entire period. To cover the entire period consistently, the second mentions are not analysed below. Third, the question indicates the relative rank of heroes. We know the popularity of an individual vis-à-vis others and can chart how the hero's popularity changes over time relative to that of others, but we do not know the absolute level of admiration and how it has changed. Finally, the contemporary focus naturally eliminates such enduring historical figures as the Founding Fathers. By forcing people to select from present-day figures we naturally have their attention directed to contemporary events.

Most Admired Man: Overall

Domestic political leaders are by far the most popular heroes for Americans, accounting for 45 percent of all mentions over the years.[1] Presidents are the most popular political figure, with incumbent presidents accounting for 18.8 percent and ex-presidents for another 7.8 percent. Presidents are followed by elected officials (15.3 percent), especially senators and governors who were presidential aspirants (e.g., Adlai Stevenson and Ted Kennedy). Judges and other appointed officials make up the remaining 3.1 percent. Following domestic political leaders are religious leaders (12.9 percent). The leading religious leaders have been Billy Graham and the Catholic popes. Third come foreign political leaders (7.9 percent), including government leaders, royalty, and UN officials. For many years Winston Churchill dominated this category, and Anwar Sadat was a major figure in the 1970s. Fourth, military leaders make up 7.4 percent of selections. Dwight D. Eisenhower (before his elevation to the presidency) and MacArthur were the leading lights. Fifth, those in science, medicine, and education account for 2.5 percent. Even this rather low rating exaggerates the level of admiration of intellectuals since the science category includes many scientific explorers (e.g., the astronauts) and few Nobel laureates. Similarly, the medical category is dominated by Albert Schweitzer—as noted for his religious works as his health contributions. Cultural figures, running the gamut from Leonard Bernstein to Elvis Presley, account for 2.4 percent. Reformers and activists, primarily civil right leaders, had 2.1 percent. That leaves at the bottom business and labor leaders (0.9 percent) and members of the fourth estate (0.8 percent). In addition, 4.9 percent selected a relative or personal friend, and 13.2 percent fell in a miscellaneous group, mostly unidentified.

---

[1] In this and other figures, the Don't know/none categories have been excluded. Tables with the Don't know/none categories excluded are available from the author.

What does this tell us about America? The clearest message is that when Americans think of admired men, they overwhelmingly think of their president and other political leaders. While much has been written about the cynicism of Americans toward politics and their distaste for politicians, Americans show a great deal of admiration for their elected leaders. Equally striking is the message we don't hear. While America is commonly pictured as a materialistic society and as the bastion of capitalists, Americans almost never pick business leaders as objects of admiration. Old money, new money, financial and manufacturing innovators, heads of multinational corporations—none get more than a rare nod of acknowledgment. In fact, the two men who make up the *majority* of business selections (Bernard Baruch and Lee Iacocca) were probably chosen as much for their public service as for their economic contributions.

Most Admired Man: Trends

Objects of admiration have shown decided shifts over the years. While the year-to-year peaks and valleys are of interest, we will restrict ourselves to examining only the general, long-term shifts. Domestic politics has shown a loss of allure over the years. If we treat Eisenhower as a political figure even before his election as president, the proportion selecting elected officials shows a decline of almost 3–4 percent per annum. Troughs were reached in 1974 after the resignation and pardon of Richard Nixon and in 1980 after Ronald Reagan's victory over Jimmy Carter. Within this general decline, incumbent presidents show a special pattern. For each president admiration waned over the course of his term in office. Yet even in the admiration ditch of Watergate, people did not fully turn away from government, for Henry Kissinger rose tremendously in esteem in 1973 and 1974 and almost matched the drop among elected politicians.

Also going down was the military. Exact trends depend on just how and when we convert Ike from five-star general to president, but military figures dwindle in the 1960s and disappear after 1977. The final category showing an overall decline is foreign political leaders. The drop in this case is in large part accounted for by the retirement and eventual death of Churchill. Both declines represent the dying echoes of World War II as the great military and political leaders of war pass on.

The two rising categories are religious leaders and personal intimates. Both increase their popularity despite well-known sociological theorizing about the growth of modern society that is secular and cosmopolitan, with weakening ties to the old institutions of community, church, and family. The rise in religion is partly a positive response to the evangelical travels of Pope John Paul II, who is much more popular than his predecessors.

While other changes may be occurring in the remaining categories (economics, culture, mass media, and science/medicine/education), the small numbers involved compared to the large numbers in the unidentified miscellaneous category make the situation uncertain.

Most Admired Woman: Overall

Over the last 30 years a near majority of Americans (48.5 percent) have selected domestic political figures as the most admired women. Former First Ladies, especially Eleanor Roosevelt and Jacqueline Kennedy, are the most popular (32.4 percent), followed by incumbent First Ladies (6.6 percent). Wives of the other major politicians as well as women holding office in their own right make up 5.7 percent, while office holders and their wives account for 3.8 percent. Foreign political leaders constitute the second largest category (15.4 percent). This consists of female heads of state and government such as Margaret Thatcher, Indira Gandhi, Golda Meir, and to a lesser extent, members of royalty—especially the British monarchy. Those in the category of culture come next with 6.1 percent, a diverse blend of both "serious" and "popular" artists. This is followed by those in the category of religion with 2.7 percent (dominated by Mother Theresa) and science, medicine, and education with 2.1 percent (mostly Sister Kenney). The remaining general categories all gather few mentions (mass media 1.6 percent, reformers and activists 1.1 percent, military 0.1 percent, and economics 0.1 percent). In addition, friends and relatives garner 7.9 percent of the choices.

Most Admired Woman: Trends

Domestic political leaders have declined as an object of admiration. This drop is particularly acute among former First Ladies. Roosevelt generally dominated the lists in the 1940s to the early 1960s, and Kennedy played a similar role until her remarriage to Aristotle Onassis, but since then no former or current First Lady has scored so high. The space vacated by political figures and their wives has been filled primarily by a gain in foreign political figures. The selection of women to head the governments of India, Israel, and England in the 1970s and/

or 1980s gave women more foreign models to draw on, and since 1971 Meir or Thatcher has led in 7 of 12 surveys.

Also showing some gain over time have been the categories of religion (primarily due to Mother Theresa's prominence since 1979) and of relatives and friends.

Man and Woman: A Comparison

Over the years it has been harder for respondents to think of a woman whom they admire than a man. Until recently, the percentage saying "none" or "don't know" has generally been 10–15 percentage points higher for women than for men. (The two exceptions were in December 1963, when President Kennedy's widow attracted such strong sympathy that there was less nonresponse about women than men, and in 1969, when a coding pecularity appears to have inflated item nonresponse for men). The margin has fallen in recent years, however, to around 5 percentage points, and in 1984 the woman question had lower nonresponse that men. This seems to have resulted from Geraldine Ferraro's vice-presidential candidacy and may or may not mark a lasting development. The general trend, however, is quite clear and pronounced. As women have filled more domestic and foreign positions of power, it has become easier for respondents to find suitable admiration models.

Perhaps the most striking similarity between men and women is the high dominance of politics and government among both sexes. Americans tend to turn first to their presidents for heroic models. Even when the presidential exemplar has been less than praiseworthy (e.g., Nixon in 1973–1974), Americans do not stray far, selecting other political and goverment leaders in the president's stead. Likewise, since it has not been possible to select a president as most admired woman, Americans have turned to the best approximations they can find, First Ladies and foreign leaders.

Differences between admired men and women result largely from sex segregation in social roles. Men lead women in the three categories where men hold almost all of the top leadership roles: the military (men 7.4 percent vs. women 0.1 percent), religion (12.9 percent vs. 4.0 percent), and science/medicine/education (2.5 percent vs. 2.1 percent). On the other hand, more women than men are drawn from the arts and entertainment (men 2.4 percent vs. women 6.1 percent), mass media (0.8 percent vs. 1.6 percent), and relatives and friends (4.9 percent vs. 7.9 percent). While the mass media category is small, it illustrates a key point. More women from the media are chosen than men though men have always held the top media positions (e.g., evening anchor, White

House reporter, editorial commentator, newspaper editors, etc.). Women lead men not because they actually dominate the field, but because their role is relatively greater (and more visible) than the role they play in such alternative fields as the military, religion, or science.

Vox Populi

Democratic idealism, not materialism, guides the admiration of the American people. Their heroes are presidential: incumbent presidents, ex-presidents, presidential aspirants, and the wives, mothers, and daughters of presidents. In turn, Americans have a tendency to try to turn their heroes into presidents. Eisenhower epitomized this tendency, and we see other examples in the political career of John Glenn and the recent draft Iacocca movement. Yet a series of failed and disappointing presidencies have diminished the public faith in politics and government. There is talk that the personal popularity and successes of Ronald Reagan have turned things around. While public admiration has risen from the dark days of Watergate and the frustrations of the Carter years, there has been no rebound to the heady years of Eisenhower and the early New Frontier and Great Society years, when government figures predominated.

When not selecting presidential and government exemplars, Americans have tended to select other people who are serving the public good or some higher power. In the early years, the military commanders of World War II were common choices. They were praiseworthy not only for their military successes, but also because they had fought to preserve and protect the country. As the cohort of World War II military leaders passed into history and the failures of Vietnam replaced the triumphs of World War II, military leaders disappeared from the admiration lists. Another group dedicated to selfless service, religious leaders, expanded into the void left by the declining political and vanishing military figures. If faith was lost in political and military power, it was gained in the agents of the Almighty.

In contrast, economic leaders have been ignored as admiration models. While Americans generally favor our free enterprise system, acknowledge that capitalists play an important role in development, and believe that private strivings lead to general gain, they perhaps believe that economic leaders receive their just rewards from the marketplace and do not venerate them. Similarly, figures in the arts and entertainment fields do not rank high. In part, this may reflect a lack of appreciation for high culture and the fine arts (most of the persons mentioned certainly fall in the popular category), but it also results from the fact that these stars are not seen as dedicated to serving the common good.

Most Admired Man, 1948–1984 (Percent)

	1948	*1949*	*1950*	*1951*	*1952*	*1953*	*1954*	*1955*	*1957*	*1958*
Political/gov. (domestic)	34.6	25.6	17.6	17.5	42.4	44.4	45.9	38.3	35.3	36.9
President	20.9	12.4	4.5	8.3	5.1	29.1	27.2	26.1	27.2	22.3
Ex-president	3.3	3.1	3.9	1.8	1.4	5.9	5.5	5.4	4.0	5.0
Other politicans	9.7	9.6	8.5	6.5	34.6	8.5	12.2	5.2	3.1	8.0
Gov. officals	0.7	0.5	0.7	0.9	1.3	0.9	1.0	1.6	1.0	1.6
Political/gov. (foreign)	5.8	13.1	8.1	6.8	5.8	10.3	12.9	11.6	7.3	11.6
Foreign leaders	5.3	12.8	7.5	6.5	5.1	10.2	12.8	11.4	6.7	11.2
UN officials	0.0	0.0	0.4	0.1	0.3	0.0	0.0	0.0	0.6	0.3
Royalty	0.5	0.3	0.2	0.2	0.4	0.1	0.1	0.2	0.0	0.1
Economic	1.0	3.3	2.5	0.9	1.6	1.4	2.5	1.4	0.0	1.4
Business	0.6	2.0	2.4	0.8	1.1	1.4	1.9	1.1	0.0	1.4
Labor	0.4	1.3	0.1	0.1	0.5	0.0	0.6	0.3	0.0	0.0
Entertainment/culture	2.6	2.0	1.7	2.0	3.1	1.3	1.7	1.3	0.0	0.4
Sports	0.5	0.7	0.7	0.8	1.1	0.2	0.1	0.3	0.0	0.0
Arts	2.1	1.3	1.0	1.2	2.0	1.1	1.6	1.0	0.0	0.4
Science/education	1.5	2.1	1.5	1.8	2.4	2.8	3.5	5.1	7.2	7.0
Military	25.6	18.9	40.8	39.7	11.6	7.5	3.2	4.0	2.8	2.8
Religion	3.1	4.5	3.6	4.2	6.3	5.9	9.8	8.4	7.5	8.9
Mass media	1.0	2.9	1.3	1.1	1.3	0.3	0.6	0.7	0.0	0.0
Activists/reformers	0.0	0.0	2.0	0.5	0.7	0.1	0.2	0.7	0.0	0.9
Relatives/friends	1.0	1.3	0.8	4.9	1.3	0.1	0.7	0.6	0.0	1.4
Misc., listed	0.8	0.3	0.4	0.3	1.08	0.1	1.5	0.4	0.0	0.0
Misc., unspecified	1.3	1.1	0.1	0.0	0.6	5.5	0.0	0.2	14.2	11.4
DK/none	21.8	24.9	19.5	20.3	22.3	20.3	17.5	27.1	25.7	17.4
	(1572)	(1497)	(1369)	(1317)	(1424)	(1483)	(1488)	(1433)	(1527)	(1509)

	1961	*1963*	*1965*	*1966*	*1967*	*1969*	*1970*	*1971*	*1972*	*1973*
Political/gov. (domestic)	45.1	40.5	43.9	48.0	40.8	33.3	29.9	29.6	32.3	34.1
President	22.9	20.7	19.1	14.0	7.5	15.7	10.0	9.2	11.4	6.2
Ex-president	17.6	13.3	12.6	12.8	13.2	5.9	3.4	3.6	6.4	0.1
Other politicans	4.4	6.2	10.4	19.0	18.2	11.2	15.3	16.3	10.6	16.3
Gov. officals	0.2	0.3	1.8	2.2	1.9	0.5	1.2	0.5	3.9	11.5
Political/gov. (foreign)	10.2	7.5	1.7	4.7	3.0	1.2	1.1	2.8	2.1	0.6
Foreign leaders	10.2	7.3	1.1	1.3	2.1	1.1	0.5	1.7	2.1	0.6
UN officials	0.0	0.2	0.6	3.4	0.8	0.0	0.3	1.1	0.0	0.0
Royalty	0.0	0.0	0.0	0.0	0.1	0.1	0.3	0.0	0.0	0.0
Economic	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.4	0.1	0.0
Business	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0
Labor	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0
Entertainment/culture	0.9	0.9	2.4	0.6	1.6	1.1	1.2	3.1	1.7	1.2
Sports	0.1	0.0	0.1	0.1	0.3	0.0	0.0	0.0	0.0	0.1
Arts	0.8	0.9	2.3	0.5	1.3	1.1	1.2	3.1	1.7	1.1
Science/education	5.9	4.1	1.8	0.3	1.4	0.8	0.8	0.4	0.6	0.3
Military	1.7	1.1	0.0	0.5	1.4	0.4	0.6	0.5	0.7	1.0
Religion	6.1	5.4	9.1	10.1	11.0	9.6	11.1	12.5	11.5	9.6
Mass media	0.2	0.6	0.2	0.2	0.1	0.7	0.5	0.7	0.3	0.3
Activists/reformers	1.1	2.5	2.8	1.5	1.5	0.1	0.8	2.4	1.3	2.2
Relatives/friends	0.2	4.7	4.0	3.3	2.3	2.6	3.2	3.9	5.5	5.7
Misc., listed	0.0	0.0	0.0	0.0	0.0	0.3	0.2	0.5	0.6	0.3
Misc., unspecified	4.7	9.5	11.9	5.2	5.9	7.6	13.5	8.9	11.6	12.8
DK/none	23.3	23.1	21.8	25.7	30.9	42.4	37.2	34.4	31.5	32.0
	(1602)	(1613)	(1565)	(1469)	(1549)	(1537)	(1546)	(1504)	(1504)	(1536)

Most Admired Man, 1948–1984 (Percent) (*continued*)

	1974	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*
Political/gov. (domestic)	34.3	24.0	24.0	27.6	15.0	29.7	26.2	27.8	22.0
President	3.0	9.5	9.3	8.2	3.7	19.7	14.7	19.6	11.5
Ex-president	2.3	3.5	6.5	2.8	3.3	5.0	5.0	2.6	2.8
Other politicans	12.3	8.4	6.9	14.2	5.9	3.2	5.4	4.1	5.8
Gov. officals	16.7	2.6	1.3	2.4	2.1	1.8	1.1	1.5	1.9
Political/gov. (foreign)	0.8	11.9	9.4	5.6	7.7	1.7	1.5	1.1	0.3
Foreign leaders	0.8	11.9	9.4	5.6	7.7	1.7	1.5	1.1	0.3
UN officials	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Royalty	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Economic	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1
Business	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1
Labor	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Entertainment/culture	1.4	2.8	3.9	1.1	1.8	2.3	2.3	1.6	1.9
Sports	0.1	0.7	1.4	0.3	0.5	0.4	0.1	0.1	0.1
Arts	1.3	2.1	2.5	0.8	1.3	1.9	2.2	1.5	1.8
Science/education	0.2	0.0	0.0	0.0	0.0	0.0	0.3	1.3	0.2
Military	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Religion	9.3	7.6	17.1	16.7	12.5	11.0	13.7	11.8	12.1
Mass media	0.2	0.7	0.3	0.3	0.5	0.5	0.1	0.8	0.5
Activists/reformers	1.5	0.9	0.8	0.8	1.5	1.1	2.1	7.2	6.2
Relatives/friends	4.0	4.7	3.1	5.2	8.0	8.8	10.3	6.2	5.0
Misc., listed	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Misc., unspecified	12.0	15.0	16.4	12.3	14.1	19.7	16.1	18.1	17.5
DK/none	36.1	32.2	25.1	30.5	38.9	25.3	27.4	24.2	32.3
	(1517)	(1518)	(1563)	(1549)	(1549)	(1508)	(1507)	(1510)	(1518)

Most Admired Woman, 1947–1984 (Percent)

	1947	*1948*	*1949*	*1951*	*1952*	*1953*	*1954*	*1955*	*1957*	*1958*
Political/gov. (domestic)	36.0	35.3	37.5	36.9	39.9	41.0	43.4	37.2	37.1	34.9
President family	1.4	1.2	0.9	1.1	1.0	4.1	5.8	4.4	3.0	4.0
Ex-presidential family	29.6	28.0	32.0	28.9	30.0	30.0	26.9	24.7	29.9	25.8
Other political	1.3	2.7	1.8	4.3	4.1	0.7	4.8	0.9	0.9	1.0
Gov. officals	3.7	3.4	2.8	2.6	4.8	6.2	5.9	7.2	3.3	4.1
Political/gov. (foreign)	10.8	13.4	3.5	4.8	6.1	11.4	5.8	6.7	7.3	7.6
Foreign leaders/wives	6.7	9.7	1.6	2.9	1.4	1.6	1.5	1.5	0.9	2.3
UN officials	0.1	0.0	0.2	0.4	0.3	5.1	1.0	0.9	0.6	0.2
Royalty	4.0	3.7	1.7	1.5	4.4	4.7	3.3	4.3	5.8	5.1
Economic	0.0	0.1	0.0	0.0	0.2	0.3	0.2	0.0	0.0	0.0
Business	0.0	0.1	0.0	0.0	0.2	0.3	0.2	0.0	0.0	0.0
Labor	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Entertainment/culture	5.0	5.2	4.6	3.9	6.8	2.0	8.2	1.2	2.9	5.2
Sports	0.1	0.1	0.0	0.0	0.2	0.0	0.1	0.0	0.0	0.0
Arts	4.9	5.1	4.6	3.9	6.6	2.0	8.1	1.2	2.9	5.2
Science/education	10.1	4.4	3.8	5.0	3.7	0.3	0.3	0.0	0.0	0.0
Military	0.3	0.3	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0
Religion	0.3	0.1	0.6	0.4	0.3	0.2	0.6	0.0	0.0	0.0
Mass media	1.6	0.8	0.5	1.7	0.8	0.4	0.4	0.0	0.0	0.5
Activists/reformers	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
Relatives/friends	3.1	3.1	3.3	2.4	2.7	0.7	1.9	2.3	0.0	2.8
Misc., listed	3.8	3.4	2.5	2.0	3.2	2.8	2.3	3.7	2.4	3.6
Misc., unspecified	0.2	0.9	1.1	1.2	1.1	5.0	0.8	5.6	9.8	8.4
DK/none	28.7	32.8	42.5	41.6	35.1	35.9	36.0	43.3	40.5	36.8
	(1482)	(1462)	(1373)	(1369)	(1424)	(1483)	(1483)	(1433)	(1527)	(1509)

Most Admired Woman, 1947–1984 (Percent) (*continued*)

	1961	*1963*	*1965*	*1966*	*1969*	*1970*	*1971*	*1972*	*1973*	*1974*
Political/gov. (domestic)	42.6	60.0	45.2	32.9	26.7	24.7	24.6	21.7	18.3	18.8
President family	5.4	1.1	3.8	3.8	3.4	4.6	4.9	9.0	4.1	4.3
Ex-presidential family	33.5	55.5	39.8	25.7	16.7	14.4	12.4	7.7	7.7	8.9
Other political	1.4	2.2	0.6	2.6	5.8	4.1	5.0	4.5	5.9	4.5
Gov. officals	2.3	1.2	1.0	0.8	0.8	1.6	2.3	0.5	0.6	1.1
Political/gov. (foreign)	5.1	2.9	5.3	10.3	13.7	8.9	15.4	12.4	19.9	11.8
Foreign leaders/wives	1.1	0.7	2.1	6.5	10.5	7.5	14.0	11.3	18.8	11.1
UN officials	0.4	0.4	0.1	1.1	0.0	0.2	0.1	0.0	0.0	0.0
Royalty	3.6	1.8	3.1	2.7	3.2	1.2	1.3	1.1	1.1	0.7
Economic	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0
Business	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0
Labor	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Entertainment/culture	1.9	2.0	5.3	2.3	5.6	2.7	4.3	4.1	3.4	3.8
Sports	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.7
Arts	1.8	2.0	5.3	2.3	5.6	2.7	4.3	4.1	2.6	3.1
Science/education	0.1	0.0	0.3	0.6	1.3	0.6	0.8	1.2	1.1	2.2
Military	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Religion	0.1	0.0	0.2	0.1	0.5	0.0	0.0	0.1	0.0	0.1
Mass media	1.2	0.5	0.3	0.5	1.2	0.6	1.1	0.4	1.2	0.8
Activists/reformers	0.0	0.0	0.0	0.0	3.1	1.7	1.5	1.2	1.4	2.4
Relatives/friends	2.5	4.9	5.8	5.6	4.2	4.7	6.0	7.0	7.7	5.9
Misc., listed	2.3	1.0	2.1	1.1	0.3	0.0	0.0	0.0	0.0	0.0
Misc., unspecified	4.1	6.3	7.1	4.2	2.7	8.0	6.1	9.2	8.3	8.1
DK/none	40.2	22.4	28.3	42.4	40.7	48.0	40.4	42.7	38.7	46.3
	(1602)	(1613)	(1565)	(1469)	(1537)	(1546)	(1504)	(1504)	(1536)	(1517)

	1977	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*1984*
Political/gov. (domestic)	21.4	27.7	20.5	15.1	19.9	12.9	13.9	26.9
President family	5.5	6.9	9.9	4.5	4.8	3.5	5.4	3.6
Ex-presidential family	10.2	15.7	8.0	4.7	9.0	5.4	5.1	3.8
Other political	5.7	5.1	2.6	5.9	1.8	2.6	2.1	15.2
Gov. officals	0.0	0.0	0.0	0.0	4.3	1.4	1.3	4.3
Political/gov. (foreign)	8.8	9.2	6.4	7.8	8.9	13.8	16.1	17.9
Foreign leaders/wives	6.9	7.7	5.2	7.0	7.2	10.4	13.7	15.9
UN officials	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Royalty	1.9	1.5	1.2	0.8	1.7	3.4	2.4	2.0
Economic	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Business	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Labor	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Entertainment/culture	5.3	3.1	1.8	2.9	4.0	5.3	2.6	3.1
Sports	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Arts	5.3	3.1	1.8	2.9	4.0	5.3	2.6	3.1
Science/education	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
Military	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Religion	0.3	0.1	4.6	6.2	5.8	8.8	10.2	7.7
Mass media	2.3	1.9	2.5	1.2	1.5	2.0	1.8	0.6
Activists/reformers	0.7	1.0	1.2	0.9	0.7	1.2	0.9	1.4
Relatives/friends	6.6	5.2	6.2	9.4	9.3	10.3	9.0	5.1
Misc., listed	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Misc., unspecified	15.2	17.4	13.2	11.3	19.9	13.7	16.2	6.8
DK/none	38.5	34.4	43.6	45.1	30.0	32.1	29.4	30.3
	(1518)	(1563)	(1549)	(1549)	(1508)	(1507)	(1510)	(1518)

Appendix: Question Wordings

	AIPO Study Number
1. What person, living today, that you have heard or read about do you admire most?	370, 371T
2. What person, living today in any part of the world, that you have heard or read about do you admire the most?	371K, 390, 425K[c]
3. What man/woman, living today in any part of the world, that you have heard or read about, do you admire the most?	408[a], 425T[c], 433K[b]
4. What woman, living in any part of the world, that you have heard or read about do you admire the most?	433T
5. What man/woman that you have heard or read about, living today in any part of the world, do you admire the most?	450TSP, 469[a], 509, 524, 540, 557, 576[c], 593, 608, 622[c]
5A. And who would be your second choice?	639[c], 653, 666[c], 681, 703[c], 721, 738, 755[b], 772[c], 793, 820, 842, 861, 885, 920, 990, 117G, 144G, 166G, 186G, 206G, 228G, 246G
6. What man that you have heard about, living today in any part of the world, do you admire the most?	469, 483
7. Which of these famous women comes closest to your ideal of what you, yourself, would like to be?	483
8. Which woman that you have heard or read about, living today in any part of the world, do you admire the most?	703
9. Here is a list of 25 prominent women, which one would you say you admire most? Is there any woman not on this list whom you admire most in the world today?	755
10. Here is a list of 20 women, which one of them would you say you admire the most? And who would be your second choice?	773
11. Here is a list of prominent women. Would you tell me which three of these women you admire the most? Is there anyone not on this list whom you particularly admire?	964

[a] Asked only about woman.

[b] Asked only about man.

[c] Data unavailable.

Occurrences of Most Admired Question

		Wordings		
Study	*Date*	*Man*	*Woman*	*Person*
A370	4/46	—	—	1
A371	5/46	—	—	1,2
A390	2/47	—	—	2
A408	11/47	—	3	—
A425	8/48	—	3	—
A433	11-12/48	3	4	—
A450	11-12/49	5	5	—
A469	1/51	6	5	—
A483	12/51	6	7	—
A509	12/52	5	5	—
A524	12/53	5	5	—
A540	12/54	5	5	—
A557	12/55	5	5	—
A576	12/56	5	5	—
A593	1/58	5	5	—
A608	12/58	5	5	—
A622	12/59	5	5	—
A639	12/60	5A	5A	—
A653	12/61	5A	5A	—
A666	12/62	5A	5A	—
A681	12/63	5A	5A	—
A703	12/64	5A	8	—
A721	12/65	5A	5A	—
A738	12/66	5A	5A	—
A755	12/67	5A	9	—
A772	12/68	5A	?	—
A773	1/69	—	10	—
A793	11/69	5A	5A	—
A820	12/70	5A	5A	—
A842	12/71	5A	5A	—
A861	12/72	5A	5A	—
A885	12/73	5A	5A	—
A920	12/74	5A	5A	—
A964	12/76	—	11	—
A990	12/77	5A	5A	—
A117G	11/78	5A	5A	—
A144G	11/79	5A	5A	—
A166G	12/80	5A	5A	—
A186G	11/81	5A	5A	—
A206G	12/82	5A	5A	—
A228G	12/83	5A	5A	—
A246G	11-12/85	5A	5A	—